PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Louis Jason Slimak

Entitled
Consilient Cognitive Literary Studies

For the degree of   Doctor of Philosophy

Is approved by the final examining committee:

William J. Palmer Co-Chair   Ryan Schneider Co-Chair
_____   _____
                Chair
Paula Leverage   Jeff Karpicke
_____   _____

_____   _____

_____   _____

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): William J. Palmer
                                _____

                                Ryan Schneider
                                _____

Approved by: Nancy Peterson   09/14/2011
             Head of the Graduate Program   Date

Graduate School Form 20
(Revised 9/10)

PURDUE UNIVERSITY
GRADUATE SCHOOL

Research Integrity and Copyright Disclaimer

Title of Thesis/Dissertation:

Consilient Cognitive Literary Studies

For the degree of   Doctor of Philosophy

I certify that in the preparation of this thesis, I have observed the provisions of *Purdue University Executive Memorandum No. C-22,* September 6, 1991, *Policy on Integrity in Research.**

Further, I certify that this work is free of plagiarism and all materials appearing in this thesis/dissertation have been properly quoted and attributed.

I certify that all copyrighted material incorporated into this thesis/dissertation is in compliance with the United States' copyright law and that I have received written permission from the copyright owners for my use of their work, which is beyond the scope of the law. I agree to indemnify and save harmless Purdue University from any and all claims that may be asserted or that may arise from any copyright violation.

Louis Jason Slimak
_____
Printed Name and Signature of Candidate

09/14/2011
_____
Date (month/day/year)

*Located at http://www.purdue.edu/policies/pages/teach_res_outreach/c_22.html

CONSILIENT COGNITIVE LITERARY STUDIES

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Louis J. Slimak

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2012

Purdue University

West Lafayette, Indiana

UMI Number: 3545409

UMI

Dissertation Publishing

UMI 3545409

ProQuest

TABLE OF CONTENTS

ABSTRACT

Slimak, Louis J. Ph.D., Purdue University, May 2012. Consilient Cognitive Literary Studies. Major Professors: William Palmer and Ryan Schneider.

Literary studies, as they exist currently, is a tripartite entity. The majority of scholarship produced is literary interpretation of specific texts. Second to interpretation is the production of "theory," those theoretical paradigms which help guide specific interpretations. There is also scholarship that concerns itself with the empirical investigation of the relationship between reader and text. Not only must these three areas of literary studies become more closely integrated, but they must also become consilient with contemporary knowledge being produced in the sciences, particularly the psychological and cognitive sciences. Interpretation of contemporary authors like Ian McEwan and Richard Powers becomes enriched by an engagement with current cognitive neuroscience, literary theory, like reader response theory and reception theory, becomes hypothesis driven and responsible to developments in the psychology of memory and discourse processing, and the empirical work already being done in literary studies is brought into contact with other disciplines where it can be understood in a wider context. Lastly, consilient cognitive literary studies have direct implications on pedagogical approaches to literature.

INTRODUCTION

Early in his book, *What Science Offers the Humanities*, Edward Slingerland describes what it's been like for him as a scholar operating within both the sciences and the humanities. Upon taking his new position at the University of British Columbia, he was allowed to create his own title, and decided upon "Canada Research Chair of Chinese Thought and Embodied Cognition." Slingerland humorously relates that while the "first half is fairly straight-forward . . . the second half usually takes some explaining" (11). What is both interesting and relevant to the discussion in this dissertation is the different reactions Slingerland experiences depending upon to whom he is speaking. At one end of the spectrum, Slingerland writes about how a colleague in Psychology chuckled at his title at his welcoming party. She (his colleague) "expressed amusement at the second half of my new job description, thinking it an oxymoron. 'Isn't all cognition embodied?'" (11). Slingerland then goes on to the almost antithetical reactions from his colleagues in the humanities for whom "the body is ultimately nothing more than an inert tabula rasa to be 'inscribed' by culture or a passive victim of power structures created by disembodied discourses" (12).

Slingerland's experiences strike a note of familiarity with me. In my career as a graduate student, I have studied under professors in a number of different departments: in the humanities: English, Foreign Languages and Literatures, and Philosophy; in the

sciences: Psychology, Biology, and Statistics. Inevitably, someone, another graduate student in one of my classes, or the professor, will ask me what I'm doing there, or, more to the point, what it is exactly that I'm studying. Unlike Slingerland, I find myself without the luxury of creating a name for the burgeoning field I'm trying to enter, and most often respond with something along the lines of "Consilient Cognitive Literary Studies," or "Quantitative Cognitive Literary Theory," or even just "Scientific Literary Theory." And, like Slingerland, depending upon which class I'm sitting in, I tend to get one of two distinct reactions.

The first, and by far the easiest to respond to, is what my humanist colleagues almost universally ask me: "Do you really think science can tell us anything new about literature, or the arts?" I'm always quick to answer to that question, yes, I do think science and, along with it, its statistical, empirical, methodological tool kit, can reveal quite a bit that we don't know about literature. Jonathan Gottschall, another scientifically-minded literary scholar, gave two simple reasons to bring a scientific approach to bear on literature:

> First, and most obviously, quantitative analysis may help us do a better job of seeing patterns in complex literary works, or large populations of literary works, that might otherwise have been overlooked or underappreciated. . . . Second, the addition of a quantitative dimension to literary scholarship will substantially improve the power and precision of strictly qualitative work." (52-53)

While Gottschall's argument is extremely general, it is also hard to deny its simple logic. Part of the inspiration behind this project is to demonstrate some of the particular benefits a more scientific methodology might bring to literary studies. While both Gottschall's and

Slingerland's books were published in 2008, there have been a number of other recent calls for science and the humanities to come into closer contact with each other, and the argument against a separation of scientific and humanistic concerns dates as far back in the Western tradition as the ancient Greeks[1].

What is new to the discussion as it has developed in the last thirty years or so is the addition of consilience to the mix. Consilience is the optimistic, idealistic legacy of the Enlightenment, a dream of epistemological unity spanning across the sciences and humanities, with knowledge claims produced in one field held responsible to related and relevant knowledge claims made in other fields. Consilience is a vehicle to intellectual solidarity and progress, driven to this end with science as its methodological engine. Consilience is also a concept that is the subject of much debate, both within the sciences and within the humanities. The argument for consilience is far more contentious than the argument for a scientific approach to literary studies, and, as such, will receive its own defense at length. But, regardless of how one feels about consilience, the potential for scientific contributions to the understanding of literature seems almost boundless. Insights from evolutionary psychology and biology allow researchers like Jonathan Gottschall, Joseph Carroll, Brian Boyd, Dennis Dutton, and David Barash to explore an evolutionary account of the development of literature, an understanding of the value of literature itself as a practice, both historically and in contemporary culture, as well as specific practices within literature, such as genres, motifs, and thematic elements.

Work done in cognitive linguistics, particularly concerning the cognitive importance of metaphor and image schema, has given literary scholars a powerful new tool for textual interpretation. The collected discoveries in the brain sciences are a major

impetus behind the so-called "cognitive revolution" in literary studies, with work being done in any number of related areas, like Lisa Zunshine's *Why We Read Fiction* which made use of the concept of Theory of Mind, Dan Lloyd's work on the narrative structure of embodied action routines, Norman Holland's career of tying psychoanalytic literary theory to "the new knowledge of how the mind works in perception, memory, learning, bilateralization, and most important for a literary critic, in the way we use language" (16).

Nearest to my own interests is the work of cognitive psychologists studying discourse processes like Richard Gerrig, Walter Kintsch, and Art Graesser, as well as cognitive narratologists like Peter Dixon, Marisa Bortolussi, David Herman, Patrick Colm Hogan, Kitty Klein, Manfred Jahn, Marie-Laure Ryan, Monika Fludernik, and Uri Margolin, all of whom are experimentally and empirically studying how readers construct, interact with, and are effected by, narrative worlds. For example, Marie-Laure Ryan has done fascinating work with reader's spatial construction of narrative space, in which she had subjects literally sketch maps representing either physical places, character movements, or "storyspace" (plot) events taken from Gabriel Garcia Marquez's *Chronicle of a Death Foretold,* in order to see the relation between the mapping type and memory for the text. In quite a different kind of study, Richard Gerrig has extensively studied the "real-world effects of fictional information," or, in other words, why people who watched *Jaws* were so reluctant to go back in the ocean afterwards (197). With their combined research projects ranging from readers' construction of spatial arrangements in narrative, empathic and ethical responses to texts, the complex interaction of textual features, emotion, and memory, and the effects of narrative construction, cohesion and causality on trauma survivors, this burgeoning approach to literature has a wealth of untapped

potential. Each of these areas is only in its infancy, with a great deal of work to be done to fully develop our understanding of the interaction of readers with narrative texts in each of these areas. And this is but a scratch on the surface of the potential for cognitive literary studies. My own particular interest is along an, as of yet, undeveloped avenue of exploration: the effects of background knowledge upon individual readers' interpretations of texts.

Despite the fact the greatest concentration of my professional study has been within the English Department, it's the second question that my colleagues in the sciences ask me that has proved much more daunting to answer, particularly in a succinct, lucid manner: "What exactly *is* literary theory?" My trouble in answering this question arises out of what I feel are the two components that would make up a respectable definition: 1) the object of study in literary theory or literary studies, and, 2) the methodology by which we proceed to study that object. The problem is that the study of literature is still what Thomas Kuhn would call pre-paradigmatic and suffers from all the resultant complications. Literary studies and literary scholars, if the field is to be empircized, scientized, quantified, or, more generally, rigorized, must realize that the discipline is still in its infancy, what Thomas Kuhn would have called a pre-paradigmatic stage. Only fully developed scientific paradigms provide answers to basic questions that guide the educational and pragmatic background from which research in that field is conducted, questions like: "What are the fundamental entities of which the universe [or text, or cognitive process] is composed? How do these interact with each other and with the senses? What questions may legitimately be asked about such entities and what techniques employed in seeking solutions?" (Kuhn 5). The problem facing a pre-

paradigmatic "science" like literary theory is neatly summed up by Jonathan Gottschall: "In contrast to the gradual, halting, yet undeniable progress of scientific knowledge, literary scholars rarely produce knowledge that can withstand the critiques of the next generation. Literary study is not, in the main, a progressive discipline where the space of possible explanation is gradually narrowed" (xi-xii). In the last century alone, there were no fewer than *ten* interpretative paradigms that simply followed one another in temporal succession, as each new generation of critics and interpretations simply displaced the last: from traditional historical and biographical criticism in the early century, to New Criticism, and then to Structuralism, Post-Structuralism and Deconstruction, New Historicism, Reader Response theories, Cultural Studies (in which I include Feminist, Queer, Race, and Post-Colonial theories), Postmodernism, with Marxist and psychoanalytic criticism enduring from the early part of the century through the current era of scholarship, even if both have been substantially reworked. But as Gottschall and a growing number of scientifically minded literary scholars are claiming, the study of literature *can* change, but *only* by employing the methods of science, for, after all, "[s]cience *is* the method" (Gottschall 12).

To return to my embarrassing failure to provide my scientific colleagues with a concise answer, a quick survey of the professional literature will neatly demonstrate my dilemma by quickly revealing not only the diverse objects of study, but also the lack of a coherent methodology in literary studies. *Contemporary Literature* is The University of Wisconsin's reputable academic journal. In its Summer 2009 issue it published the following articles: 1) an interview with poet Rae Armantrout; 2) an article entitled "The Closures of the Open Text: Lyn Hejinian's 'Paradise Found,'" described in the article's

introduction as an exploration of how author Lyn Hejinian uses *"openness* to signify an emphasis on linguistic opacity, autonomy, and polysemy and the rejection of organicist notions of naturalness, presence, and immediacy . . . [and] exhibits a striking preoccupation with total linguistic transparency, correspondence between language and world, epistemological closure, and perfect understanding, all of which she associates with the term *paradise"* (Edmond 240); 3) an article on author J.M. Coetzee's *Slow Man*, which is, in part, a response to critic Derrik Attridge's critical work on the author, and in part a study of the Derridean trope of hospitality in *Slow Man*; 4) an examination of the poetry of and correspondences between Robert Duncan and Denise Levertov for gendered representations of social space; 5) a description of the use of musical tuning as a metaphor for political exile in Daniel Mason's *The Piano Tuner* and Vikram Seth's *An Equal Music*; 6) "Scratching the Threshold: Textual Sound and Political Form in Toni Cade Bambara's *The Salt Eaters*," which explores the link between language use and racial politics through a novel; 7) four brief reviews of recent fictional and non-fictional works.

Within this *single issue sample* there are no fewer than seven particular objects of study: 1) an author; 2) a particular literary (poetic, non-fictional, or fictional) text; 3) the corpus of a single author's work; 4) linguistic structures within fiction; 5) literary criticism carried out by other critics; 5) gender issues; 6) political issues; 7) racial issues. Even if we were to condense this list so that objects 1 through 3 became a new object 1) literary authors and their works, object 4 remained unique (as new object 2), and objects 5 through 7 became new object 3) socio-political issues, there are still at least three objects of study in this single issue of a single literary journal, and I feel quite

comfortable in claiming that a more diverse sample would only add to this list. Literary authors and their works have been the traditional focus of literary studies – the appreciation, elucidation, and interpretation of literary works produced by those authors deemed worthy. Linguistic and semiological approaches to literature are representative of an earlier trend in literary studies to become more quantitative and scientific, but this approach often proceeded without science's empirical methodology. However, since the 1960's and the rise of what is alternatively called Postmodernism, Post-Structuralism, Liberationalism, or just plain old Theory, it is the "study" of socio-political issues as existing withing texts that has been the dominant domain of literary studies.

So, then, if we were to use this small sample of literary criticism to answer my scientific colleagues as to what the object of study in literary studies is, it would have to be either: *Object 1: literary works and the authors that produced them; Object 2: the linguistic and semiotic structures and practices embodied in literary texts; Object 3: engagements with contemporary and historical socio-political issues as exemplified in literary texts; or, Object 4: a combination of any two or all three of these objects.* While diverse, these objects of study seem perfectly natural to a field whose primary object of study must always include a literary text. Literature, by its very definition, is the representation of life through texts, whether prosaic or poetic. As any art-form can, then, literature is capable of engaging with any and every other aspect of life, science, religion, thought, or experience. Thus literary studies and scholars are bound to come into contact with almost infinitely diverse phenomena.

How then does literary studies address the analysis of these near-infinitely complex phenomena, or, in other words, what methodological constraints are practiced in order to ensure reputable and respectable scholarship, and to contribute to the progression of knowledge in the field? Again, I will use the single issue survey of *Contemporary Literature* to answer this question, but with drastically different results this time. The interview is based on an actual dialog performed by the interviewer and the author, while the articles all share a single methodological approach: dialogic engagement with other, mostly literary, dialogic arguments. If in terms of an object of study literary studies is diverse and complex, its methodology is paradoxically antiquated and unevolved: dialogic authoritarianism[2]. This isn't to say that there isn't value in this sort of methodology, only that, as so many empirical, scientific, quantitative, and/or consilient-minded scholars have recently noted, that the scientific method is the best system yet devised by mankind for producing a durable, cumulative, knowledge base that can be systematically improved upon. As Jonathan Gottschall so astutely notes, since the field has undeniably migrated towards "an intense absorption with the cacophonous social, political, and economic contexts in which texts are produced and consumed . . . [i]t matters whether we are right or wrong in the strong claims of fact we regularly make about the nature of gender, sexuality, human competitive tendencies, ethnocentrism, language, [and] oppression" (78-79).

Literary studies is not the first discipline to find itself in a pre-paradigmatic crisis, nor the first whose practitioners feel that the scientific method can't possibly be applied to their subject matter, that the diverse phenomena represented in the complex cognitive interactions between reader and text simply cannot be quantified, or as Andrew Delbanco

so glibly claims, that literary studies will quite simply "*never* be able to submit its hypotheses to the scientific test of replicable results" (37; emphasis added). Theodore Porter's book, *The Rise of Statistical Thinking*, provides one striking and surprising analagous situation to contemporary literary studies: the practice of medicine in the 18th and 19th centuries. Porter details how early attempts to introduce statistical quantification and scientific methodology into medicine was met with almost universal opposition, with opponents claiming that the scientific approach to medicine would fail to account for "the variability of medical facts, which could only be fully appreciated through induction and medical intuition" (Porter 159). Another historian of statistics, Stephen Stigler writes that "[i]n the 300 years since Newton's *Principia*, mathematical probability and statistics have found application in all the sciences – social, physical and biological. In each area where these ideas have been introduced there has been resistance as the protectors of the different realms have sought to prevent the 'Queen of the Sciences' [mathematics and statistics] from conquering new territory" (203).

Not only is it hard to imagine medical science, let alone biology, any of the social sciences, or the physical sciences, as *unscientific* or *unquantitative* disciplines, it is almost impossible to imagine feeling comfortable visiting a contemporary doctor who openly rejected the findings and methods of medical science, or to purchase a car engineered without concern for physical laws, or to feel that political policies that ignored the findings of the social sciences were sound. Yet those same disciplines were once pre-paradigmatic themselves, with similar doubts and reservations about making use of scientific methods as literary scholars do today. It is my perspective, however, that the fears and misgivings of resistant literary scholars should instead be replaced with hope;

the hope that a turn to consilient, quantitative, empirical methods in literary studies will result in the same boon that has resulted in *every* other scientific discipline, and reveal previously unknown patterns and processes in literature, as well as deepen our already profound aesthetic appreciation of literary works[3,4].

At this point, I feel the need to define a few of my terms, so that both my readers from the sciences and the humanities understand me. I'm going to start with the least problematic of the terms which is "empirical." By empirical literary studies, I mean what most of the sciences mean: a procedure for studying a subject rooted in experimental control, manipulation, observation, replication, and statistical significance. Studies conducted by Je'meljan Hakemulder on the empathic, ethical, and moral effects of literature, Marisa Bortolussi's and Peter Dixon's narratologically-focused, psychologically-methodological investigations, and Peter Gerrig's experiments examining the discourse processing of literary texts, are all quality examples of this line of work. "Quantitative" literary studies are those which make use of statistical analysis to provide evidence for their claims, with the studies being done in evolutionary literary studies by Joseph Carroll and Jonathan Gottschall as sound representatives.

Which leaves "consilience". Defining the word "consilience" is, itself, a simple task. Literally, consilience is a "jumping together," typically taken to mean a "jumping together" of disparate facts and phenomena under a common theory. Charles Darwin's theory of natural selection and the evolution of species is perhaps the quintessential example of this kind of consilience, where disparate theories and data from anthropology, paleontology, and any number of particular fields within biology were united under one theoretical edifice. However, since being coined by William Whewell in 1840, the term

has been the center of a semantic and philosophic dispute governing the proper scope of its potential application. As there are (at least) two camps with distinctly different views on the scope, possibility and even desirability of achieving universal consilience across the disciplines, I feel it is necessary that I give each position a brief explanation before I declare where I stand and why.

Sociobiologist E.O. Wilson single-handedly revived the term in 1998 with his full length book on consilience, aptly entitled *Consilience*. Wilson describes his view of consilience as a "metaphysical world view" in which literally all knowledge produced through scientific methodology within every discipline is brought into harmony (9). Wilson views the study of all phenomena - and he specifically targets the resistance by those in the humanities who study the phenomena of "human action" (i.e. literature, and the arts in general) and the historical "unfolding of unique events" - as governed by "physical causation," and thus argues that if "ten thousand humanoid histories could be traced on ten thousand Earthlike planets, and from a comparative study of those histories empirical tests and principles evolved, historiography – the explanation of historical trends – would already be a natural science" (11). Wilson's idealistic epistemological view is essentially concerned with the complex intersections of what are now methodologically diverse fields separated *not by fundamental differences in subject matter but, rather, by what he calls* "artifacts of scholarship," *or methodologies and ideologies* (9). For example, while there is work being done in each of the respective fields of ethics, biology, social science, and environmental policy, Wilson claims that there is alarmingly little work being done that makes use of all knowledge in all four disciplines at once in order to produce truly consilient knowledge about the ethical social

interactions between human beings and our environment, and how best to legally govern those interactions.

Wilson's call for unification across the gap that separates the humanities from the sciences has resonated with a large number of scientists and humanist scholars alike. Humberto Maturana and Francisco Varela's *The Tree of Knowledge* is essentially a Wilsonian argument for consilience via "the biological roots of understanding" (11). Evolutionary psychologists Leda Cosmides and John Tooby have written about "vertical integration," another reworking of Wilsonian consilience with their additional clarification of a non-hierarchical, topographical relationship, with the simplest phenomena and most concrete scientific laws (i.e. physics, chemistry, and biology) at the "bottom" of the "tree," moving upwards into more complex phenomena and less reductive, more descriptive explanations (i.e. psychology, the social sciences, and the humanities). Humanist scholars Jonathan Gottschall and Edward Slingerland are likewise Wilsonian in their belief that by practicing the methodology of science, the humanities will not only benefit by seeing their disciplines become cumulative and progressive, but related disciplines that are now separated by the methodological gap (i.e. discourse processing in psychology and narratology in literary studies) will be brought into closer contact with one another, and thus mutually benefit one another through the sharing of relevant knowledge. While Harold Bloom's "anxiety of influence" *may* describe the relationship of those who *write* literature from one generation to the next – and even this is an empirical question – it should hardly be the case for those who *study* that literature.

The take home message from all of these scholars is that knowledge in one field shouldn't contradict knowledge that has been produced in another. When there is conflict,

such as there is between literary psychoanalytic models of thinking and current psychological models, it needs to be resolved, with preference given to where the majority of empirical evidence lies. To those who would argue that this *seems* to put humanistic scholarship at a disadvantage because positions within the sciences have a much more firmly established quantitative and empirical tradition, I would say the humanities *are* at a disadvantage. Consilience in our scholarship is a method to remedy that disadvantage. This isn't a position that makes the humanities subordinate to the sciences, or one that makes the sciences invulnerable to change brought about by theories produced in the humanities, only a claim that radical differences in knowledge claims need to be resolved; hardly a revolutionary position. To return to my example of Freudian and psychoanalytic models of cognition in literary studies, while the significance of Freud's work for Modernist writing is *historically* important for its *influence* upon those writers, Freud's model has since been rejected by the natural sciences, and thus fails to explain the cognitive processes *actually* used in writing, reading, and interpreting those texts, or of the characters who are engaged in "thinking" within the narrative worlds. Put another way, while it may be perfectly sound to examine Modernist texts *for evidence of Freudian structures within those texts*, it is neither consilient nor prudent to then construct *models of writing, reading, and literary interpretation upon those same Freudian* (or Lacanian, et al.) *models* which would be in direct conflict with contemporary models generated by the psychological sciences. Instead of this conflict, literary theorists practicing consilient cognitive literary criticism would find themselves uniquely placed to enter into a discourse with their scientific colleagues surrounding the

cognitive practices of reading, writing, the perceptual processes of semantic recognition, memory studies, emotional and social cognition, or linguistic processes.

Consilient literary critics could also then produce relevant traditional literary criticism concerned with representations of consciousness, embodiment, space, and the post-human or cybernetic future, to name just a few possible avenues of consilient criticism. Of particular interest to me is post-humanism, that thread of philosophy which examines the interfaces between biological organisms and their machinic tools, of which cognitive artifacts are of particular interest. A dominant trend in literary post-humanism is an unfortunate reversion to Cartesian dualism, an antiquated philosophical perspective that maintained distinct differences in existence between the physical and the mental. Consilience with current trends in embodied cognition both within biology and cognitive psychology would not only highlight post-human inconsistencies with the sciences, but also draw attention to what are quasi-fictional reinterpretations of the literature under consideration. For example, speaking of cyberpunk fiction, the fictional genre with the closest thematic ties to post-human thought, Sheryl Vint writes that cyberpunk is "a genre best known for its rejection of embodiment and embrace of an existence in cyberspace" (*Bodies* 103). Vint goes on to show her Cartesian loyalties when she proclaims that "cyberspace is the consummate world of the Cartesian dualist: in cyberspace one *is* the mind, effortlessly moving beyond the limitations of the human body" (*Bodies* 103). Similarly, Vicky Kirby explains that cyberspace is "the space where the perfect body is paradoxically acquired through an annihilation of the flesh" (132). In William Gibson's *Neuromancer,* for example, not only does the novel primarily take place within the real world as opposed to cyberspace, but *every* single access of cyberspace is described in

terms of the protagonist's physical, embodied interface. Therefore, not only do I feel that unconsilient post-human interpretations are misreadings of the fiction which they claim to represent, but, worse yet, they are dramatically inconsistent with the emerging field of embodied cognition which claims that the mind (human or otherwise) and the body are one dissoluble unit, incapable of being understood apart from each other.

However, there is another, less optimistic view about consilience between the sciences and the humanities, one put forward in its mildest view from evolutionary biologist Stephen Jay Gould, and in its most extreme form by literary scholar Louis Menand. While Gould favors the sort of interdisciplinarity urged by the Wilsonian camp, he also argues that the sciences and the humanities are two distinct magisteria, and that nothing science does will ever fully breach the gap. Gould has two keen criticisms of the universal consilience project: 1) he argues that certain systems, particularly complex biological systems, are emergent and contingent; 2) he claims that the proper goal of science is to explain what *is*, and the role of moral reasoning (the only concrete example he gives of a humanistic study that is definitively *not* a study of what is) is to explain what we *ought* to do. For me, it is only the first of Gould's two objections that merits any response; his second claim is essentially Kantian in its character, and the complex relationship between neuroscience, biology, evolution, and ethics has an extensive history all of its own that is beyond the scope of this project to even summarize, not to mention that plenty of disciplines within the humanities would argue that they do study what "is" within their respective fields.

But Gould's first criticism rests on two fundamental assumptions, which may, in fact, turn out to be correct: the first is that emergence is, in fact, an actual property of life and not a theoretical construct used to cover up current areas where our understanding is incomplete, and the second, which is conditional upon the first, is that contingent occurrences cannot be predicted. Gould defines emergent properties as "properties that make their first appearance in a complex system as a consequence of nonadditive interactions among components of the system," which is a technical way of saying that in certain systems, like human culture or cognition, there may be properties of those systems, like literature or consciousness, that emerge inexplicably from the sum of the components that produce them, and no reductive (i.e. scientific) analysis of those components could *possibly* predict the emergent properties (223). While the issue is far more complex than this, my reaction is that those who claim emergent properties can *never* be predicted or explained are using science and mathematics to construct a new dualism that still preserves some of the classic metaphysical structures as somehow irreducibly unique and special: the mind instead of the soul, emergent life instead of the miracle of creation, etc. While the complexity of systems that *seem* to display emergent properties may be overwhelming now, I trust in the method and the progress of human knowledge to eventually discover the principles that give rise to emergent properties. My response to his argument for historical and individual contingency is precisely the same. In any case, the only way to settle the argument with Gould is to wait and see, and continue with our intellectual endeavors in the meantime.

While Gould favors consilience across disciplines where it is applicable, he merely doubts that it will be possible to achieve universal consilience, with all

phenomena in all disciplines equally explained by one set of integrated principles. Literary scholar Louis Menand, on the other hand, views consilience as "a bargain with the devil" (14). In his article for *Profession*, the Modern Language Association's (literary studies' governing body) journal concerned explicitly with the profession of literary study, Menand bemoans the current state of the discipline, and claims that what the "profession could use [is] some younger people who think that the grownups got it all wrong" with post-everything theory, but, for reasons he never explicitly lays out, that any "young Turk" claiming that consilience and scientific approaches to literature is the way to go is absolutely wrong (12). Menand even goes so far as to suggest that beyond simply rejecting consilience, humanities departments should also avoid true interdisciplinarity, and instead "hunt down the disciplines whose subject matter they covet and bring them into their own realm" (14). Most humanities departments already practice this sort of interdisciplinarity; not by actually studying in other related fields, but by "hunting down" useful bits of jargon or productive ideas and co-opting them beyond their intended meaning. However, this is exactly one of the problems with the field as a whole, the very problem that inspired Gottschall's and Slingerland's books, as well as this dissertation, all of which are considered responses to Menand's near-incomprehensible program. With respect to this dissertation, however, Menand's suggestion also highlights the problem within the more focused approach of cognitive literary studies, which I will address a in Chapter One. H. Porter Abbott says cognitive literary studies, for all its novelty, must avoid being merely "an approach . . .manned by a bunch of scholar-pirates who plunder for their purposes troves of hypotheses, bright ideas, and . . . rigorous scientific work" (714). The best way to avoid this sort of selective importation of jargon is the use of a

transparent methodology, like that of science and the consilient method for which I am arguing here, which facilitates the evaluation of the claims made within the operative paradigm.

One of Menand's closing quips perfectly underscores the problem with the sort of intellectual seclusion that currently reigns in the humanities. Menand writes: "What Derrida believed about how language works is not what the average newspaper reporter believes about how language works. Why is that a scandal? What are philosophers for? For that matter, what are universities for?" (16). What Menand is missing *is* the problem, and the problem is that Derrida had theoretical *beliefs* about how language worked, beliefs that were never empirically tested or verified, and that those beliefs were in conflict with the routine *knowledge* that an average newspaper reporter has about how language actually effects people. The problem with a non-consilient approach is that this is an accepted situation, that it *isn't* a scandal that people who use language on a daily basis and know what it can do, and people who theorize about the nature of language don't have a common ground for testing their hypotheses and coming to an understanding that grows out of a scientific methodology. Let me be clear: it *is* a scandal. It is *the* scandal that threatens the humanities. Universities are for the generation of knowledge, cumulative, lasting knowledge produced, not at whimsy, but through rigorous testing and refinement. The humanities, particularly the field of literary studies, haven't been playing by the rules and that is why they find themselves in crisis. The old guard of high theory *did* get it all wrong, and they got it all wrong not because of *what* they said, but because of *how* they said it, because they shunned scientific methodology, and they didn't care that their theories came into conflict with what other fields had already shown to be false.

Consilience isn't a bargain with the devil, it's a path back to relevance, respectability, and the production of scholarship that will be durable, meaningful, cumulative, and progressive.

I am closest in spirit to being a member of the Wilsonian camp, and this project has been heavily influenced by the assumptions that go with that set of beliefs. If postmodernism has taught us anything valuable, and I believe that it has, it has taught us to acknowledge our biases, and to likewise acknowledge that those biases impact our work in fundamental ways. Where I depart from Gottschall and Slingerland is that I feel that they are too narrow in their application of consilience to literary studies. Gottschall claims a consilient literary framework should be developed under evolutionary biology's umbrella (while sometimes not staying true to that methodological approach), while Slingerland makes the same claim for embodied cognition. I feel that a consilient approach should be taken *in general*, with specific effort made to bring a particular type of study into contact with all of the relevant disciplines, humanist and scientific. New Historicist critics should study historiographic methodology and be aware of the trends, theories, and data in that field, evolutionary critics like Gottschall and Carroll are beholden to evolutionary biology, cognitive critics to psychology and the neurosciences, and social critics of all kinds (feminist, gendered, sexualized, racialized, colonialized), must make use of what is being studied and done in biology, psychology, and the social sciences, including sociology and anthropology. As Jonathan Gottschall writes, "A consilient perspective encourages and endorses historical, biographical, linguistic, economic, philosophical, psychological, anthropological, biological, and sociological

approaches to literary study," but without "flout[ing] great swaths of theory and data" from the respective sciences (38).

As with the multiple avenues of research that Gottschall outlines for consilient evolutionary literary studies in his book, I feel that a consilient cognitive approach to literature also has a number of potential directions to apply to interpretation and around which to develop research programs. The most important of these are:

1) Mirror neuron research which influences conceptual formulations of Theory of Mind or intersubjectivity, as well as empathy. This can be related to perspective taking in literature, and developing empathy for characters, as well as formulating ethical judgments of those characters.

2) Affective and emotional processing in the brain and its evolutionary development. This can help identify textual features which give rise to the complex emotions aroused by literature.

3) Recursive models of consciousness posit reentrant (also sometimes called recursive, reverberant, or recurrent) and backward-building neuronal structures, which potentially explain how the brain "'clos[es] the loop' between past and present activity, or between predicted and actual versions of the input" (Tononi & Koch 250). In other words, recursive models are responsible for creating what Gerald Edelman calls the "remembered present," a present moment in consciousness that draws upon the immediate experiential past, as well as the mnemonic past, as well as looks forward towards the expected future. It would seem highly relevant to explore the relationship between these theoretical structures of

consciousness to the use of metafictional forms in contemporary fiction, which are themselves recursive.

4) Perhaps most importantly, as Edward Slingerland himself has argued, the idea of embodied cognition. This holds that our brain is embodied in our body and our body is embedded in its particular econiche, and that the two have an inseparable evolutionary relationship, with dramatic consequences for normal cognition, and even language use, particularly in metaphorical constructions. This theoretical assumption comes into direct contact (and conflict) with a great deal of posthuman literary scholarship, and is also directly relevant to a richer understanding of a great deal of contemporary cyberpunk and science fiction. Moreover, it may also have dramatic implications for the actual cognitive processing of texts depending upon their physical forms; i.e., are there affective, cognitive, and/or mnemonic differences between reading *Hamlet* from a leatherbound folio edition, a cheap Penguin paperback, or an on-line hypertext?

Consilience, for me (and for the rest of its use in this dissertation), is a concept that demands agreement and fidelity in theories and vocabularies cutting across disciplines and fields of study, without a hierarchical priority given to a particular field in settling disagreements, but with priority given to where the bulk of the empirical evidence lies. Knowledge produced in one field, be it in the humanities or in the sciences, that blatantly contradicts knowledge produced in another field, calls attention to a conflict that needs to be resolved, preferably through interdisciplinary cooperation between the involved fields.

A call for consilient literary theory is equivalent to a call for paradigmatic consolidation within the field, or, in other words, a call for a change in focus from strictly producing literary interpretations and theories to including the empirical study of the processes of literary interpretation themselves. This isn't to say there isn't space for literary interpretation, only that textual interpretations and elucidations are themselves interesting *if and only if* we can comprehend those interpretations within a larger framework of cognitive processes that could potentially account not only for aggregate patterns in interpretative communities, but also explain some of the tremendously complex and varied individual responses to particular texts. I agree wholeheartedly with Gottschall's and Slingerland's explanations of what this means for humanist and literary scholars: cross-listed courses with the sciences and other relevant disciplines, required courses in statistical analysis and empirical and quantitative methodologies, and cooperative studies and publications with members of different departments, not just "the cherry-picking of confirming evidence and importation of jargon that have often passed for interdisciplinarity" (Gottschall 75). However, what both Gottschall and Slingerland fail to address is one of the major concerns of what consilient studies mean to *individual* scholars versus *disciplines* and *interpretative paradigms* as a whole. For example, a cognitive literary critic doesn't necessarily have to be familiar with all of cognitive science in order to practice consilient cognitive literary criticism, so long as the paradigm of cognitive literary criticism has been developed in a consilient manner. One simple move towards this end is to have the institutional practice of peer-reviewed journals accommodate a structure whereby scholars from the humanities and sciences could

review relevant claims being made within each others' respective fields, as well as to encourage more inter-departmental studies and publications.

Like Gottschall and Marcus Nordland, I also recognize the perils for such a fledgling interdisciplinary program where there are "no negative tests" established yet, and where scientific theories can be chosen "on the basis of ideological preferences rather than the criterion of truth," and the "widespread assumption that theories can simply be transformed willfully whenever they do not meet the requirements of the interpreter" (Nordland 313). To this end, consilient literary theory needs to address the theoretical basis of a general consilient methodology for cognitive and scientific literary studies. The work of empirical literary scholars like David Miall, Willie Van Peer, Janmeljan Hakemulder, Richard Gerrig, Maria Bortolussi and Peter Dixon, Jonathan Gottschall, among a growing number of others, has already lain the foundation for empirical literary studies to build upon. However, as most of these theorists are well aware, it is also important for consilient literary theorists to address and be aware of the ideological and pedagogical implications of such an approach, as well as to define ways for literary critics to generate credible, testable hypotheses and literary interpretations. This dissertation is a direct descendant of the work of these kinds of literary scholars, and therefore I feel it necessary to specify what that means. I will not be arguing any further for the *need* for consilience with the sciences, as I believe that argument has been effectively made in a number of other places. Either you feel that consilience is desirable, or, like Louis Menand, you don't. Nor will I address a *basic* empirical methodology, as this too has been effectively covered in a number of sources[5].

What I will do can be broken down into three basic components: In Chapter 1, I specifically address what consilient cognitive literary criticism is, what it isn't, what is relevant to its development in the fields of literary studies, linguistics, psychology, and the cognitive neurosciences, where it came from, and where it's going. If, to this point, my readers in traditional literature departments may be feeling left out in the cold, or, perhaps more realistically, unsure of how a consilient approach to literature does anything more than turn the study of literature into a branch of psychology (or any of the other relevant sciences), then Chapters 2 and 3 should bring them in from the cold, to paraphrase John LeCarre. These two chapters will demonstrate how consilient cognitive literary criticism can engage with traditional literary criticism and interpretation by acting as a guide to its interpretations of literary texts being produced within the current scientific paradigms.

The claims of humanities scholars that indict consilience do so variously for its "conflation, simplism, ontological reductionism, [and] scientism," as well as out of a "Romantic belief in the autonomy of the human spirit, a postmodernist epistemological nihilism, and a traditional humanistic belief in the irreducible singularity of all artistic productions"(Wilson 11; Carroll 395). Yet, an appreciation for art is inseparable from understanding the social, political, and historical context in which it was produced, including the dominant scientific paradigms. Much as the works of Modernist writers were heavily influenced by the psychological theories of William James, Sigmund Freud, and Henri Bergson and the formulation of relativistic physics by Albert Einstein, Niels Bohr, and Werner Heisenberg, late 20th and 21st century literary texts are clearly beholden to the findings and theoretical positions of contemporary cognitive science and physics.

Thus, another goal of consilient cognitive literary criticism is to produce interpretations of literary texts from the viewpoint of the psychological paradigm that is contemporary to its production; while it may make sense to look for Freudian influence in Modernist fiction, it makes far less sense to look for Freudian structures in late twentieth century texts, let alone to use Freudian formulations that have been discarded by the natural sciences to explain literary texts almost a century removed from the period of his influence.

The potential for this approach to literary interpretation extends well beyond the contemporary fiction upon which I will be focusing. Indeed, as Alan Richardson has lamented:

> Few critics have as yet produced cognitively informed interpretative readings of literary texts that at the same time fully acknowledge their historical specificity. . . . [I]ssues in literary history, far from being occluded by approaches that recognize the validity of human universals and species-specific cognitive mechanisms, can be productively reopened in ways that have eluded criticism that relies on purely constructivist notions of the subject. (5)

In Chapters 2 and 3, I engage contemporary works of fiction, like *Saturday* and *Enduring Love* by Ian McEwan, the Vietnam War fiction of Tim O'Brien, Phillip K. Dick's *Do Androids Dream of Electric Sheep,* and William *Neuromancer*, and the depictions of the mind therein with related theories and findings from the relevant cognitive sciences.

An increasing number of critically acclaimed authors like Ian McEwan and Richard Powers who are making explicit use of cognitive neuroscience in their novels, and for the basis of their construction of represented consciousness. Powers and McEwan

are among the more vocal about the influence of neuroscience on their writing; both men often reference leading neuroscientists like Gerald Edelman, V.S. Ramachandran, and Antonio Damasio in their interviews, and in the dedication to McEwan's *Saturday,* the first person named is "Neil Kitchen, MD, FRCS (SN), Consultant Neurosurgeon and Associate Clinical Director [of] The National Hospital for Neurology and Neurosurgery" in London, whom McEwan observed and consulted with for two years prior to writing the novel. McEwan's novel, *Saturday*, provides a concise, yet nuanced example of what I will do at length in Chapters 2 and 3.

The novel begins with a scene where shortly after Henry Perowne, the novel's protagonist neurosurgeon, wakes in the middle of the night, he finds himself standing at the window, looking at the sky, before he sees something that is initially unidentifiable. McEwan creates Henry's progressive awareness of the stimulus by representing his changing visual perceptions of the event and then linking those discriminations to the feelings that accompany them. McEwan's model of non-representational, predictive and interactive visual perception is best understood within the recent work on Bayesian sensory perception by Martin Banks, Daniel Kersten, Richard Gregory, and J.J. Gibson, among a host of others. In this model, sensory perception, in this case visual perception, occurs not as a veridical representation of what is "out there" in the world, but begins as a model of what the brain expects based on evolutionary constraints as well as personal experiences. When Perowne first sees the streaking light, there is little precedent or context with which to make sense of the sight, and thus he undergoes a quick succession of changes in perception, which, quite like the model of consciousness put forward by Daniel Dennett, may or may not enter awareness as "multiple drafts." Moreover, the link

in the scene between Perowne's perceptions and the resultant feelings and emotions finds itself most fully expressed in the work of neuroscientist Antonio Damasio. In what is a four or five page scene, McEwan's fiction directly engages with at least three major areas of contemporary psychological and neuroscientific research.

To refresh the reader's memory then, the novel begins with Henry Perowne becoming aware of a flash of light in the night sky, when, "in his eagerness and his curiosity, he assumes proportions on a planetary scale: it's a meteor burning out in the London sky" (McEwan 12). However, Perowne quickly apprehends that this object is "moving slowly, majestically even," and, so, four sentences later, in that indeterminate amount of time between his original content-fixation of the flash, and then his flawed perception and awareness, meaning is recreated somewhere in what Daniel Dennett calls the intentional loop, a model of consciousness involving "multiple drafts" of what is being perceived (McEwan 12). J.J. Gibson's theory of affordances in which the shape of objects is inseparable from their semantic meaning and motor utility, as well as the Bayesian models of prediction and error correction are also on display here. The end result is that Perowne's brain quickly has to revise his first draft of the situation. Notice that I say his brain and not Perowne himself. Key to Dennett's idea of multiple drafts, Gerald Edelman's reentrant circuits, Damasio's second order representations, and the overwhelming majority of models of consciousness in cognitive neuroscience, is the belief that the brain is not equivalent to the mind, although the mind is entirely dependent upon the activity of the brain, and that activity is recursive, constantly in contact with and reacting to its former states. In other words, in this scene and situations like it, it is, at least, ambiguous whether or not Perowne himself is conscious of the revisions as they

take place, or, if conscious, if they are fully conscious (articulated in thought) or more felt. To return then to the streaking light, Perowne (or his brain) revises his initial judgment; it's not a meteor hundreds of miles up into the atmosphere but, rather, "a comet," "millions of miles distant, far out in space swinging in timeless orbit around the sun" (McEwan 13). He goes to wake his wife up to share in the unexpected awe of the sight when "he hears a low rumbling sound" and realizes that, one, a comet would be so far out in space that it would appear stationary in the sky, and, two, not make any noise. Though "only three or four seconds have passed since he saw this fire in the sky," he's now changed his mind for the second time, this time revising the scale of the draft back to "the local," watching as a burning plane screams through the night (McEwan 13).

It is perhaps also relevant to note the existence of multiple pathways for sensory stimuli and cognitive, semantic and emotional processing in the brain. Joseph LeDoux's work with the visual cortex and the amygdala (the area of the brain primarily responsible for fear responses) has shown that there are (at least) two separate pathways for potentially threating visual stimuli (snakes, spiders, etc.). Visual signals from the optic nerve first travel to the visual thalamus, a relay station of sorts for sensory information. From there, depending upon the potential threat level of the stimuli, the information is relayed along a longer path through the visual cortex for detailed processing, and then into the amygdala for emotionally appropriate evaluation, *and/or* directly sent from the visual thalamus straight to the amygdala, allowing for the immediate autonomic responses that are evolutionarily salient in a threating situation (increase in blood flow, elevated heartbeat, preparation for flight, etc.). In a situation like Perowne's, what McEwan could be attempting to represent is just such cognitive processing, with detailed

visual processing occurring in parallel but lagging behind the emotional evaluation. Clearly, before he ever knows exactly what it is he is seeing, Perowne is aware that it is unusual, of some importance, and, potentially, even a threat. Perowne's final observation moves from "eagerness and curiosity," to a "leap of gratitude for a glimpse, beyond the earthly frame" that is "extraordinary." It evolves into a "nightmare," a burning plane making an emergency crash landing that finally disappears from sight (12-14). These few moments of thought brilliantly capture the neuroscientific description of recursive consciousness, and of the ambiguous separation between conscious and unconscious, emotional and rational processes that give rise to the higher-order consciousness that both readers and writers most often privilege. More importantly, this brief scene, as written by McEwan *after* two years of intensive study under Neil Kitchen, bears little or no resemblance to any of the psychoanalytic structures of thought that currently dominate literary criticism. Full appreciation for McEwan's use of present tense, first-person limited narration in the novel, which only ever shifts into past tense when Perowne is explicitly remembering, and of McEwan's contemporary portrayal of consciousness can only be achieved *if* the concerned literary critic is familiar with contemporary theories of consciousness, and consilient with contemporary psychology and neuroscience.

Literary texts are incredibly complex pieces of language, and our interaction with them as readers is even more complex. Yet, despite Stephen Jay Gould's warning that these systems are contingent and emergent and thus finally escape any reductive analysis, it is only through the work of identifying the constituent components within literature, narrative, fiction, literary genres, etc., and then by addressing the effects of those components on readers in aggregate, that we can come to a foundational understanding

for the cognitive processing behind our experience of narrative worlds. In Chapter 4, I move beyond the intersections of contemporary fiction with contemporary psychology and cognitive neuroscience, and by way of a short series of empirical experiments, attempt to build upon the work of other empirical and quantitative literary scholars, and related psychological studies. Specifically, my interest is in exploring the interaction of background knowledge upon textual interpretation by way of manipulating authorial knowledge. Raymond Gibbs has begun this particular line of research in his book *Intentions in the Experience of Meaning* where he claims that authorial knowledge not only effects our interpretations of literary texts, but, once gained, is quite difficult to ignore in our interpretations, even with verbal reports to the contrary. My initial hypothesis was that authorial knowledge is most active at moments in texts that are directly relevant to the narrative world: i.e. the race, sex, or sexuality of the author in highly racialized, gendered, or sexual texts, the author's biographical history in historical texts and metafictional texts. My experiment set out to examine the effects of authorial knowledge on reader interpretation in metafictional texts, and the initial results of the experiments seem to suggest that there is indeed a predictable, statistically significant relationship between a reader's background knowledge of an author and how that same reader responds to the text. This one simple experiment has dramatic implications for the not only the comprehension of literary processing, but how we approach the pedagogy of literature.

Literary studies is a pre-paradigmatic discipline facing a crisis during a time when the lack of production of relevant knowledge capable of impacting the lives of the general public means less funding for literary departments, less social prestige which in turn

means less jobs at lower salaries resulting in a downward spiral of lower enrollment numbers, and even less general social importance. As nearly all people are consumers of stories in some form, this lack of impact upon the general public is even more dismaying. My suggestion to revive literary studies and to ensure its relevance not only to the general public but to the larger academic community as a whole follows in the tradition of other arguments for consilience from numerous other scientists and humanists. Instead of isolating ourselves, our students, and our scholarship from the larger social and academic communities with which they are concerned, we need to aggressively pursue active interdisciplinarity and the production of cumulative, progressive, consilient knowledge, in the hopes that our scholarship will make a substantial difference upon the lives of writers and readers of all kinds, interests, and expertise. To stay the course, to ignore all the warning signs and instead continue to maintain the status quo, that is the real bargain with the devil.

Notes

1)   See Jonathan Gottschall's *Literature, Science, and a New Humanities,* E.O.

Wilson's *Consilience*, Stephen Jay Gould's *The Hedgehog, the Fox, and the*

Magister's Pox, *Literary Reading* by *David Miall, The Tree of Knowledge* by

Humberto Maturana and Francisco Varela, *Psychonarratology* by Marisa

Bortolussi and Peter Dixon, and Edward Slingerland's *What Science Offers*

the Humanities.

2)   I'm by no means suggesting *all* literary scholars or scholarship is of this mode;

this is obviously *not* the case as I make repeated references here and

throughout this work to scholars and scholarship that aren't of this mold.

However, when describing something to an outsider for the first time – and I

assume most of my colleagues in the sciences to be unfamiliar with the general

state of things in literary studies – it is best to tell them that most of our birds

have feathers and fly.

3)   For more complete arguments on the need for, benefits of, and development

and practice of consilience, quantitative, and/or empirical approaches to

literature, as well as other humanistic disciplines, see: Jonathan Gottschall's

Literature, Science, and a New Humanities, E.O. Wilson's *Consilience*, Stephen

Jay Gould's *The Hedgehog, the Fox, and the Magister's Pox*, *Muses and Measures*

by Willie van Peer and Frank Hakemulder and Sonia Zyngier, *The Moral*

Laboratory by Jemeljan Hakemulder, Richard Gerrig's *Experiencing Narrative Worlds, Literary Reading* by *David Miall, The Tree of Knowledge* by Humberto Maturana and Francisco Varela, Edward Slingerland's *What Science Offers the Humanities*, and *Psychonarratology* by Marisa Bortolussi and Peter Dixon, to name only a few.

4) The misplaced fear that scientizing literary studies will somehow mar our aesthetic appreciation for literature as an art form finds an eloquent reproof from physicist Richard Feynman in an interview with the BBC in 1981:

I have a friend who's an artist and he's sometimes taken a view which I don't agree with very well. He'll hold up a flower and say, "look how beautiful it is," and I'll agree, I think. And he says, "you see, I as an artist can see how beautiful this is, but you as a scientist, oh, take this all apart and it becomes a dull thing." And I think he's kind of nutty. First of all, the beauty that he sees is available to other people and to me, too, I believe, although I might not be quite as refined aesthetically as he is. But I can appreciate the beauty of a flower. At the same time, I see much more about the flower that he sees. I could imagine the cells in there, the complicated actions inside which also have a beauty. I mean, it's not just beauty at this dimension of one centimeter: there is also beauty at a smaller dimension, the inner structure...also the processes. The fact that the colors in the flower are evolved in order to attract insects to pollinate it is interesting -- it means that insects can see the color. It adds a question -- does this aesthetic sense also exist in the lower forms that

are...why is it aesthetic, all kinds of interesting questions which a science knowledge only adds to the excitement and mystery and the awe of a flower. It only adds. I don't understand how it subtracts.

5) See Gottschall's *Science, Literature, and a New Humanities*, Miall's *Literary Reading*, Bortolussi's and Dixon's *Psychonarratology,* Gerrig's *Experiencing Narrative Worlds*, van Peer's *Muses and Measures*, and Hakemulder's *The Moral Laboratory*.

CHAPTER ONE

THE CRISIS IN LITERARY STUDIES

Part I. State of Crisis

The "crisis" in the humanities is by this time a well-documented phenomenon, at least in terms of its symptoms[1]. In 1999, Robert Weisbuch, then president of the Woodrow Wilson National Fellowship Foundation, had this to say concerning the crisis:

> Today' s consensus about the state of the humanities – it's bad, it's getting worse, and no one is doing much about it – is supported by dismal facts. The percentage of undergraduates majoring in humanities fields has been halved over the past three decades. Financing for faculty research has decreased. The salary gap between full-time scholars in the humanities and in other fields has widened, and more and more humanists are employed part time and paid ridiculously low salaries. (B4)

Weisbuch is one among many who have addressed the alleged decline in the humanities. However, few scholars have attempted to then turn an objective eye on the field and attempt to diagnose the source of its illness. Jonathan Gottschall is one of those who has addressed the causes of the crisis, and he has suggested that the problem is primarily methodological; literary scholars produce scholarship that is not consilient with contemporary knowledge produced in the sciences, most often not empirically supported,

avoids negative evidence, fails at being (or even attempting to be) politically and ideologically disinterested, and is couched in "cunningly-placed thickets of impenetrable prose" (85). In the Introduction to this work, I have already suggested that inconsilient research conducted with no regard for scientific methodology makes up at least one causal vector in the complex problem facing the humanities, and that literary scholars may begin to address the crisis by revising those practices.

The humanities find themselves in crisis not because of a failing interest in the arts, or humanistic concerns, but because of the way in which they engage those subjects. In 1990, twenty years ago now, Bernard Bergonzi made a suggestion for remedying the crisis in literary studies along these lines. Bergonzi's suggestion, developed at length in his book, *Exploding English* was "to split off conventional literary study from the activity now known as 'cultural studies,'" by which he means to separate those with scientific sensibilities towards literature from those purists who "want to read *literature*, and possess, or hope to develop 'literary sensibility'" (Kermode 618). This suggestion is nothing short of disciplinary suicide. Bergonzi's plan is the surest way to resign Literature departments to an ever decreasing role as mere cultural caretakers, and to ensure that they have no active part in the accumulation and development of knowledge about that same literature and how it is read. The most acceptable method of study that is currently available to us is the scientific method, and it is by making the study of literature more scientific in practice (as opposed to merely in its terminology) that we can begin to resolve the crisis in literary studies. It is entirely possible that the field will not embrace the necessary change, and it will find that its enrollment numbers continue to dwindle, its scholarship will be effaced, and the literature department will become a department

existing in miniature, with the exciting, relevant work concerning literature being outsourced to psychology departments and the social sciences.

While this is indeed a drastic claim calling for nothing less than a complete methodological paradigm shift within literary studies, beyond the methodological changes that must be made, this is not as drastic a revolution as it might seem to be at this point. The constellation of interpretations produced by literary critics over the last several hundred years already functions as wealth of empirical data to be studied, and even as a form of statistical aggregation; while not cumulative and progressive in the scientific sense, literary criticism has an established tradition of interpretations that essentially form a dialogic consensus. What the field needs to understand is that the production of literary interpretations is only one facet of its proper scientific activity, and most likely a minor facet at that.

In order for literary studies to become empirically grounded, it should continue the study of the cognitive bases of its production, comprehension, and interpretation, its relationship to normal cognition, its historical and evolutionary development, and its current place within contemporary cultures *through empirical and quantitative measures*. The scientific method is a potential remedy to one particular avenue of impotence within literary studies. By employing the scientific method within literary studies, we stand to not only make our scholarship more accountable, but more applicable to real life, more useful in addressing and potentially solving real world problems, and more popular because it is more accountable, applicable, and useful. However, as other scholars have so astutely noted, the production of consilient, empirically supported, disinterested, and clearly written literary theory and critical interpretations will not go far to cure the crisis

if the problem is not simply with literary methods, but with the much larger issue of whether or not literature is worth studying in the first place.

Clearly, I believe it *is* important to study literature. Literature, as a cultural practice, is thriving, despite the waning public interest in what literary critics have to say. Narrative forms proliferate in almost every aspect of public life: politics, entertainment, sports, art, and the sciences, to name only a few. Narrative-based therapy is on the rise in helping patients who have sustained psychological trauma, as well as in a range of treatments for those diagnosed under the autism umbrella, and these treatments can only be improved if we improve our understanding of the cognitive bases of and interaction with narrative. There is also a clear public interest and passion for literature and narrative within our culture, with the exponential expansion of both the publishing industries and the film and television industries. So what has happened to bring about a situation where the public no longer cares what the "experts" have to say, and, more importantly, do we even have any need for these so-called "experts"?

Ironically enough, literary theory itself may provide a hypothesis. In the 1970's, Stanley Fish generated what he called "reader-response" criticism, an approach to reading literature that was intended to focus upon the reader's engagement with the text, even if it ultimately ignored empirical engagements with actual readers. Within Fish's theoretical framework, he proposed that interpretations of literary texts were guided not solely by textual features that could be rigorously isolated and identified, but also by the cultural, social, and ideological practices of what he called "interpretative communities." While Fish never empirically tested or developed his ideas, the idea is intuitive, and seems to offer a possible explanation for what has happened to so thoroughly remove literary

theorists from the concerns and the concern of the larger world. Literary critics, particularly those practicing from within academia, have become a cloistered, near hermetic interpretative community. As Frank Kermode has noted, the "continuance of contact between experts and the educated public at large is no longer thought a plain necessity of intellectual and social health" (610). This isn't to say that it isn't desirable to have a trained interpretive community, only that we, as literary critics, have to recognize how our professional training may negatively impact our scholarship. It does so in at least two ways. The first is that literary scholars sometimes produce theories and interpretations that may have little in common with the reading practices, concerns, and interests of the public. This is the case for a small portion of all scholarship in all fields, and, as such, shouldn't be the cause of much concern. However, what is alarming is that a great deal of what we do study – race relations, gender and sexuality, theoretical ideas like Jean Baudrillard's simulacra (an idea which now proliferates in popular media, in movies like *Avatar*, books like Don DeLillo's *White Noise*, and first person video games like *Halo*), as well as narrative itself – is of great interest to the public, but what we say about it is ignored strictly on the basis of methodological considerations.

Part of the problem, as I have already suggested, is that literary studies are by and large not cumulative, with successive interpretative paradigms seeking to replace rather than refine previous paradigms. As Thomas Kuhn so elegantly explained at length in his book *The Structure of Scientific Revolutions*, while paradigm shifts are certain to occur in any field, they are bound to be the exception, not the rule. The last of Robert Weisbuch's "Six Proposals to Revive the Humanities" was this:

Embrace contradiction. While we must insist on learning for its own human sake, we also must connect the humanities to the immediate challenges in our culture. To make the world safe for private scholarship that is deliberately, grandly, rightly unconcerned with consequence, *we need to become newly public*. That means requiring students to learn how to explain their work to non-humanists. And it means that all of us must speak up. We must make the case for the value of a liberal-arts education, and for the sense that the humanities make possible the thinking about values and creativity that no technology can produce--and without which any democracy will fail. (B5; emphasis added)

Weisbuch is right in calling for literary scholars to become "newly public," and right when he claims that the public is "just bored" with "deconstruction or Marxism or whatever" (B4). How then is literary studies, so steeply grounded in interpretative paradigms like deconstruction and Marxism, to once again become relevant to the public without simply replacing deconstruction or Marxism with a newer incarnation?

Literary scholars, as trained professionals, need to stop asking themselves how they can find more "evidence" of Freudian structures in literature, or of postcolonial, racial, and sexual, political practices, and return to the generation of responsible knowledge by asking relevant questions in our research. What is it that literature does for us as individuals in our normal, day-to-day lives? What does it do for humanity as a species? How does it produce these effects? Is it, as Jemeljan Hakemulder, Wayne Booth, and Martha Nussbaum have claimed, a "moral laboratory," where we can simulate empathic and ethical interactions with other human beings? Does reading literature

improve general critical thinking, or mnemonic skills? As Richard Gerrig, Hayden White, and Paul Ricoeur have asked, can an understanding of the cognitive processing of narrative and fiction reveal something about how readers use fictional information to draw conclusions about real-world events and behaviors? Is there any affective, cognitive, or mnemonic differences for reading a hardbound leather folio edition of Shakespeare, a cheap paperback version of the text, or an on-line hypertext, a la embodied cognition? By studying the cognitive processes involved in reading, can we come to a better understanding of social comprehension that might guide pedagogical policy, and psychological treatment for sociopathological patients? These questions share two underlying characteristics. The first is that these issues are relevant beyond the walls of Literature departments, with direct implications for psychological treatment, pedagogical practices, the entertainment and marketing industries, computing technologies, as well as political studies and jurisprudence. The second trait each of these questions share is the more important for this project; these are all matters open to empirical evaluation.

Consilience, as both a methodological and theoretical framework for conducting research, is the best, currently available, single solution to the crisis in the humanities, and as such, this chapter is primarily focused on changes that need to be instituted in order to bring about consilient scholarship. These changes can be divided into two levels: 1) a general, theoretical level consilience, demanding that theories in one field don't *continue* to exist in a state of conflict with theories from another related field 2) a methodological conscilience in which methods, practices, and goals are adapted from the

general scientific method to a specific line of research, as exemplified here by cognitive literary studies.

The call for methodological changes within the humanities have received well-developed attention elsewhere, and as such will only be cursorily dealt with here[2]. The most necessary methodological changes are:

1. Institute an empirical, statistically-based, quantitative analysis as a foundational aspect of every undergraduate and graduate student's requisite curriculum within the humanities.

2. Teach the formulation of hypotheses. A great deal of literary scholarship starts because of a scholar's interest in a topic, and becomes a discursive exploration. These sorts of projects are certainly valuable for their organization of archival data, and their interpretative contributions. However, far too often they are not hypothesis driven when they just as easily could be. A hypothesis is a question that can direct research, both for the involved scholar, as well as those scholars who wish to build upon that work by refining the hypotheses, the conclusions, or challenging the validity of those conclusions.

3. As a direct consequence of the second suggestion, humanities scholars must learn to produce scholarship that can be falsified, or refuted. As Frederick Crews has said, one common factor underlying most of literary studies is the "refusal to credit one's audience with the right to challenge one's idea on dispassionate grounds" (228). We know, at an intuitive level, that all literary interpretations are not equal, no matter what the dictates of

post-structuralism may have held. We must also do away with practices within any theoretical edifice that takes evidence *against* its hypotheses as evidence *for* those same hypotheses, as does psychoanalysis, Marxist criticism, and a good deal of social criticism. Self-contained theoretical systems like these are immune to revision from within or without, and, regardless of their "explanatory power," offer little in the way of cumulative and progressive knowledge.

4. Learn to conduct controlled experiments with manipulated variables which can then be replicated.

5. As Jonathan Gottschall has argued, reverse the trend begun by Liberationist theorists, and return to what Matthew Arnold argued was the governing principle of literary criticism: disinterestedness. While post-structuralism may have shown the biases inherent in every discipline and discourse, science included, this is not a reason to give up disinterested approaches to scholarship, but a challenge to attempt to meet those biases head on, and deflect them, as much as possible, from interfering with our work.

6. Purge our prose of postmodernism's characteristic stylistic embellishments and linguistic obfuscations. Humanist professors can no longer "[invent] arcane dialects to keep out the uninitiated," but instead need to focus on tearing down "the fence between academic criticism and the intelligent world outside" (Kermode 610-611).

7. Bolster not only interdisciplinary *study*, but cooperative *publishing* both within the discipline and interdisciplinary. The sciences assume that a single study will necessarily entail the involvement of multiple people, each of whom contributes by way of their specific specialty. It is not unusual for a single neuroscientific lab to have several dozen people working jointly on a single project, from just as many angles. Kristof Koch, a leading neuroscientist at The California Institute of Technology, in visual perception, has a lab that features twenty-eight people, in fields as diverse as bio-engineering, systems administration, computational psychology, electrical engineering, computer science, electrophysiology, physics, statistics, computer engineering, neural systems, and biology, all of whom are involved in the joint publication of the lab's research. Many labs, like neuroscientist Giacomo Rizzolatti's in Italy, also employ scientifically-minded philosophers like Alvin Goldman specifically to help connect their research to on-going debates in other related discourses. In 2007, Thomson Scientific, released a survey of scientific papers that were published between 1993 and 2006. The survey found that in 2005, "[m]ore than 750 papers with 50 or more authors were published . . . Papers with more than 100 authors grew . . . to an impressive 475 . . . [and] Interestingly, papers with 500 or more authors increased from 40 in 2003 to 131 in 2005" (drugs.com). The mean number of authors per scientific paper in 2006 was 3.8. The mean number of authors per literary paper for the same year had no significant statistical difference from 1. It is a highly

questionable assumption that any individual scholar can regularly produce methodologically sound scholarship if it is to be consilient with a wide range of disciplines.

8. Imitate the faculty-mentor/student relationship within scientific graduate studies. This is the only major addition to what is an already established argument for methodological change. While mentoring is itself not methodological, it directly promotes sound methodology and research practices by giving graduate students concrete experience working on established research projects with practicing researchers.

These eight steps are just the beginning of a program to initiate a sea-change that would move the humanities away from its current decline and towards a rebirth as a more relevant cultural science, but these recommendations are general. The specifics for any particular application of these principles will depend upon the line of research undertaken: e.g., cognitive literary studies. The governing relationship between recommendations at the general level and recommendations for the specific line of research is a matter of refinement. While it is generally necessary for all consilient scholarship to be disinterested, produced in clear prose, open to refutation, and able to be replicated, studies conducted under the umbrella of cognitive literary studies, for example, must have their own clearly defined jargon, processes for falsification, and criteria for replication. While the practices of clearly defining terms, limiting neologisms, ambiguous prose, and the misappropriation of metaphor should be expected across all disciplines and lines of research, the specific jargon that must be defined in any field will naturally differ, as will the necessity for coining new words, or applying metaphorical

structures to aid in conceptual comprehension. The best way to clarify these potential differences is to lay out a plan for a consilient methodology for cognitive approaches to literary studies, and to draw attention to areas of potential differentiation with other approaches as we go along.

Perhaps the best way to explain how a consilient cognitive literary studies is to proceed is to begin by explaining cognitive theory as it exists now, and which practices need to be avoided in the future, which practices need improved upon, and which practices are already part of the foundation of a consilient approach to literary studies. Cognitive theory is a dauntingly broad term, with a history that is entangled with that of psychology itself. It could be arbitrarily dated back to include the ancient Greeks, Sigmund Freud, the cognitivism of Wilhelm Wundt, Max Wertheimer, Wolfgang Kohler, Kurt Koffka, and Jean Piaget (who were interested in the brain as a processor of information), or the cyberneticians of the 1940's and 1950's, such as Norbert Wiener, Claude Shannon, and Louis Couffignal, whose computational approach to thought and mind gave rise to the modular cognitivism of Jerry Fodor and others. The parallel history of cognitive literary theory is likewise as old and diverse as the history of psychology itself, and is nothing less than a book-length topic; there is at least one such book forthcoming[3]. For the purposes of this dissertation, I am concerned with two distinct contemporary movements: one within cognitive psychology and one within cognitive literary criticism. Within psychology, I will limit my focus to the rise of the embodied movement in cognitive psychology since the 1980's, when researchers like George Lakoff and Mark Johnson in cognitive linguistics, Rodney Brooks in AI, Gerald Edelman in cognitive neuroscience, and a host of other scientists and philosophers began to challenge

the strictly computational, implicitly dualistic approach to the mind. After examining the ideas of embodied cognition as they relate to literary studies, I will then focus on the current state of cognitive literary theory, and the course it must take in order to become consilient.

Part II. The Challenge of Embodied Cognition

The idea of embodied cognition is simple and incredibly relevant to literary scholars. The work of cognitive linguists like George Lakoff, Mark Johnson, Mark Turner, and Gilles Fauconnier has already been embraced in movements like cognitive poetics. Cognitive poetics holds that metaphorical constructions are linked to embodied image schema, and complex and abstract thoughts and representations are understood through embodied structures that give them meaning4. And while cognitive poetics implicitly embraces embodiment, that is only the first step towards consilience. As Gerald Edelman has said, "[I]t is not enough to say that the mind is embodied; one must say how" (Bright Air 15).

Edelman himself provides a clear and concise definition of embodiment as it is understood within the cognitive neurosciences:

The brain is embodied and the body is embedded. First, consider embodiment. All of the activities [of consciousness] depend on signals to the brain from the body and from the brain to the body. The brain's maps and connections are altered not only by what you sense but by how you move. In turn, the brain regulates fundamental biological functions of your body's organs in addition to controlling the motions and actions that guide your senses. These functions are the most fundamental aspects of sex, breathing, heartbeat, and so on, as

well as the responses that accompany emotion. If we include the brain as your

favorite organ, you *are* your body. Second, consider your embeddedness. Your

body is embedded and situated in a particular environment, influencing it and

being influenced by it. This set of interactions defines your econiche, as it is

called. It is well to remember that the human species evolved (along with the

brain) in a sequence of such niches. (*Second* 24-25)

Of the embodied perspective, Raymond Gibbs Jr. states that it emphasizes "the

importance of kinesthetic action in the theoretical accounts of how people perceive, learn,

think, experience emotions and consciousness, and use language" *(Embodiment* 3).

Proponents of embodied cognition do not treat the brain as a disembodied information

processor without reference to its biological (neuronal) substrate. Nor do they discount

the physical, human body, of which the brain is a vital aspect and without which it could

not survive. Additionally, the embodied perspective recognizes the importance of both the

evolutionary and contemporary econiches in which the organism operates.

Embodied cognition is more than just a psychological perspective, it is also a

philosophical response to one of the most pervasive formulations of mind-brain-body in

the Western tradition: Cartesian dualism. Rene Descartes's influence upon neuroscience

and Western thought cannot be overstated. Of Descartes's influence, neuroeconomist Paul

Glimcher writes that "[i]t is almost an axiom in scholarly circles that neuroscience as we

conceive of it today, began in the seventeenth century with the work of the French

mathematician, philosopher, and physiologist Rene Descartes" (5). Descartes's ideas are

foundational to the discourse of mind and body, so much so that, like Glimcher, other

neuroscientists like Antonio Damasio, Gerald Edelman, Shaun Gallagher, and Joseph

LeDoux, cognitive philosophers, like Daniel Dennett, John Searle and Andy Clark, and phenomenologists in the tradition of Maurice Merleau-Ponty, have devoted chapters or even entire books to undermining Cartesian dualism in its more insidious aspects (like the Dualism that is still often encountered in literary theory, which will be the subject of Chapter 2). Of Descartes's continued influence, Damasio writes, "It would not have been possible to present my side of this conversation without invoking Descartes as an emblem for a collection of ideas on body, brain, and mind that in one way or another remain influential in Western sciences and humanities" (*Descartes'* 247). Though the strong form of Cartesian dualism – a true and total separation of res cogitans, things mental, and res extensa, things physical – is widely rejected both in the sciences and humanities, Damasio is right to note that it remains unquestioned in assumptions that separate the mind from the brain. This kind of implicit dualist argument holds that the "mind and brain are related, but only in the sense that the mind is the software program run in a piece of computer hardware called brain," or will admit the existence of a relation, "but only in the sense that the former cannot survive without the life support of the latter" (*Descartes'* 248). A form of Cartesian dualism is implicit in every discussion of a human being in which the body is observed without taking into account the mind-brain with which that body interacts, regardless of whether that body is a feminized, racialized, queered, colonized, gendered, or sexualized one. The discursive practice of separating the mind-brain from body has dramatic consequences for the critical literary readings that tacitly accept that assumption, poststructural, postcolonial, feminist, queer, posthuman or otherwise.

Among the most damaging of those consequences is an acceptance of the dualist split between mind-brain and body, and thus a theoretical inability to address the reciprocal relationship between them, as well as an implicit acceptance of another faded psychological paradigm: behaviorism. Behaviorism was popularized through the work of John Watson and B.F. Skinner in the 1930's through the 1950's. Oddly, behaviorism was non-dualistic, instead treating mental phenomena like thoughts and feelings primarily as epiphenomenal and non-causal. Subjective mental states were the result of somatic states and had no relation to behavioral products, and thus were not suitable for objective, scientific study. The philosophical implications of behaviorism were clear; while it was certain that the mind-brain did something, it was irrelevant when studying a subject's behavior. Behaviorism, as a scientific paradigm, attempted to completely objectify subjects as collections of behavioral data with the goal of predicting behavioral responses from environmental stimuli.

Most structuralist and post-structuralist literary theory adopts a social constructionist perspective that apes behaviorism in its presentation of complex cognitive behavior by reducing it to its end product: social behavior. Literary critics working in this tradition are trained to look at phenomena like sexuality, race and gender performance strictly from a behaviorist or social standpoint without trying to account for the cognitive processes and structures that give rise to and are, in turn, affected by, that behavior, nor the more complex, interactions between mind-brain, body, and culture. Ironically, literature, which has historically been thought to instruct and entertain its readers through the ethical exploration of simulated experiences, has seen literary criticism leave the field of explaining what it is to be like a character, and rather, move towards a more pseudo-

scientific goal of explaining the rules of the world in which a character lives. Cognitive neuroscience, on the other hand, has shrunk from trying to explain or predict the phenomena of the exterior world and begun explaining what it is like to be an embodied, thinking, feeling human being embedded in a rich and diverse environment.

In sharp contrast to the dualist's split between mind and body are the embodied, embedded, and recursive models of consciousness offered by proponents of embodied cognition. Cognitive neuroscience has shifted the emphasis in the mind-body problem of Cartesian dualism into the mind-brain problem, where the two terms, "mind" and "brain" refer to "two different levels of explanation for the same thing, but not two different kinds of thing" (Ward 4). The most basic parts of the equation then are the mind-brain, the body itself, and the organism's econiche or environment, all of which develop together, exerting mutual, bi-directional influences upon one another. Neuroscience has primarily focused on mapping the mind-brain relationship, and it has found, unsurprisingly, that the mind-brain relationship must be one of close correlation. For instance, fMRI studies have revealed the existence of the fusiform face area, an area of the brain that shows activity when a person sees or even imagines a particular face. Therefore, the contents of the mind – the idea and identity of that face – are closely correlated to but not completely equivalent with the activity of a specific area of the brain (Frith 23). Lesion studies have confirmed again and again that damage to particular areas of the brain leads to particular changes in mental abilities. However, the relationship between mind and brain is not a perfect one-to-one. As neuroscientist Chris Frith explains, "There can be changes in the activity in my brain without any changes in my mind. On the other hand I firmly believe that there cannot be changes in my mind without

there also being changes in brain activity" (23). To reiterate what has already been said, the distinction between brain and mind may be thought of as more of a descriptive difference, than of one in kind.

The mind, or what some philosophers and neuroscientists call higher-order consciousness or self-awareness, arises out of a combination of physical structures Gerald Edelman calls reentrant or recursive pathways in the brain, mental processes that Antonio Damasio calls second-order representations, and which Bud Craig locates, at least partially, in the insular cortex, which is a key structure in the regulation of bodily homeostasis – the representation of bodily state and changes in body state over time. In essence, the brain is able to take signals and information about the body and represent the body as well as those signals themselves to itself, or to take memories and records of past experiences and make them the object of sensation or knowing themselves, which then enables consciousness of an emotion, feeling, thought, or of oneself. It is this recursive property of the mind-brain that propagates the cycle of psychosomatic representation that forms the most basic core of identity; one's body sends signals to the brain, which are made into representations of the body, which change the mind-brain's expected perception of the body, which changes the body's signals to the brain, and so on. Coupled with the brain's ability to learn via prediction, this ability to represent one's past thoughts and self to one's current mind-brain, and compare those past states and experiences to the present somatic state, physical environment, and expected results, generate the highest level of human cognitive achievement.

The embodied perspective comes with several theoretical implications. One, changes in the mind-brain *can* result in changes in the body, and changes in the body *do*

result in changes in the mind-brain. Two, because the body is embedded in the econiche (which is more than just the physical environment in which human evolution has occurred but now also includes social and cultural influences), it can bring about changes in the econiche, and the econiche, in turn, can bring about changes in the body as well as the mind-brain. Neuroscientists like the Italian team lead by Giacomo Rizzolatti who discovered mirror neurons, which in turn uncovered the link between perception, motoric simulation, learning, and performance, and even higher level cognitive processes like empathy, are a prime example of scientists concerned with the embodied aspect of consciousness. Evolutionary biologists and psychologists like Merlin Donald, who study the development of physical and cognitive structures in step with the natural and artificial alteration of the human econiche, concern themselves more with the embedded component.

If the assumption that underlies dualism is a discrete separation of *res cogito* from *res extensa*, the findings of contemporary neuroscience are revealing precisely the opposite, that the separation between mind, body, and world is really quite plastic, even within the brain itself. In a famous experiment published in 2001, Shigeru Obayashi et. al. used functional brain mapping to demonstrate that neurons in the parietal cortex would alter the effective personal space of the monkey after it had been taught to use a rake as a tool. In the experiment, Obayashi's group imaged the activity of the parietal cortex in monkeys when a piece of fruit was placed on a table in front of them and found that the neurons were only activated when the food was within physical reach, essentially a mental representation of personal-effective space. When the fruit was placed out of reach, the neurons were inactive. The monkey was then given a rake. When fruit was placed out

of arm's reach but within range of the rake, the neurons were inactive *until* the monkey received training to use the rake as a tool to get food. At that point the neurons that represented personal space were suddenly activated when the fruit was placed anywhere within reach of the rake. As Chris Frith summarized, "[a]s far as this part of the brain is concerned, the rake has become an extension of the monkey's arm" (62). In other words, the embedded body is part of the world to the mind-brain, albeit a privileged part, and the demarcation between body and world which we may take as phenomenologically static and well-defined, is quite malleable within the brain itself.

Part III. The Embodied Cognitive Approach to Literary Studies

Taken as a whole, the cluster of theoretical perspectives that make up the embodied perspective can have a direct effect upon literary studies in at least three ways:

1) Consilience with the psychological, biological, and evolutionary consequences of embodiment should directly influence literary interpretative paradigms in their theoretical formulation of the body and embodied identity. For socially minded theorists interested in the literary representation of race, gender, sexuality, or posthuman "dis-embodiment" (more on this in Chapter 3), an understanding of the psychological and biological consequences of embodiment seems necessary. Authors who foreground issues of bodily normality, authors like Flannery O'Connor, Dorothy Allison, Jeannette Winterson, as well as cyberpunk and science fiction writers like William Gibson, Neale Stephenson, and Phillip K. Dick, could be read in a revealing new light to a culture whose continued integration with informational and bio-medical

technologies is rapidly challenging our understanding of what biological, embodied normality is.

2) An understanding of embodied cognition could also bring about a fuller appreciation of contemporary portrayals of consciousness inspired by embodied models. For example, again, looking for Freudian models of consciousness in modernist works makes perfect sense because those works were produced in an era where Freud's ideas influenced the aesthetic representation of consciousness. Carrying on that practice in a time when Freud is barely more than a footnote in most psychological and neuroscientific textbooks ignores the influence that contemporary writers owe to the current scientific paradigm. Writers like Ian McEwan, Richard Powers, and Mark Haddon are explicitly drawing on contemporary neuroscientific theories of consciousness in their work. The rise of the metafictional impulse in contemporary fiction also should be understood in relation to contemporary models of consciousness that stress the recursive character of neuronal structures.

3) The first two implications of the embodied perspective primarily deal with the interpretative level of literary criticism. However, I have been stressing that the interpretative activity is actually less important than the generation of empirically supported hypotheses which engage literary processing itself. Work on literary processing carried out under the embodied paradigm could be used to develop models of textual interpretation and comprehension that are dependent upon embodied interactions and cognitive structures. In other

words, aside from focusing on the purely cognitive aspects of literary processing, such as semantics, emotional affect, text comprehension, and the textual features which give rise to these and other cognitive effects, we could also begin to ask questions that concern the physical experience of reading, and determine their relationship, if any, to literary processing. It is certainly a phenomenologically different experience to read Shakespeare from a centuries old, leather bound folio in a darkened, hushed library, seated in a plush armchair, surrounded by the smells of parchment, dust, and wood varnish, than it is to read the exact same text from a Penguin paperback seated in coach on an airplane, than it is to read an electronic copy on a small computer screen in an ill-lit dormitory room that reeks of stale beer, spoiled food, and body odor. What is unknown is how these embodied situations effect emotional, interpretative, and mnemonic performance in response to the same text, or, indeed, whether questions along these lines are worth investigating at all.

In essence, a consilient approach to an embodied understanding of literature would engage with the previous three categories each of which could be understood as falling into one of the following broad terms: the body in literature, the mind in literature, and the embodied literary experience. In order to make clear how consilience should be practiced within each realm of concern, each area will be addressed in turn.

Literary interpretative paradigms since structuralism are most adept at describing the second part of the embodied formulation: the effect of the econiche (society and culture, in particular) on the organism. Siobhan Somerville, in the introduction to her book, *Queering the Color Line*, describes a recursive process of identity formation, not

between mind-brain and body, but, rather, between action and social perception. She writes, "One's sexual identity, *while at times linked directly to one's sexual activities,* more often describes a complex ideological position, into which one is interpellated based partly on *the culture's mapping of bodies and desires and partly on one's response to that interpellation*" (6; emphasis mine). For Somerville, race, like sexual identity, "refers to a historical, ideological process rather than to fixed transhistorical or biological characterstics" (7). Or, perhaps even more simply, "one's racial identity is contingent on one's cultural and historical location" (7). Somerville's nuanced exploration of the historical and cultural effects on sexual and racial identity is praiseworthy for drawing attention to how sexuality and race are indeed at least partially figured by culture.

What Somerville's recursive structure shares with other constructivist and performativist notions of race, gender, or sexuality, is a behavioral, dualist emphasis on only one half of the recursive structure of mind-brain and body. The body does indeed interact with the socio-cultural world beyond its somatic borders, which does indeed have distinct cognitive effects upon the individual. However, the mind-brain also interacts with itself and with the body, sometimes without response or feedback from the environment, and that interaction is as important for understanding an individual's gendered, sexual, or raced identity as is the recursive social relationship. There is a growing understanding within cognitive neuroscience of body-related disorders like anorexia nervosa, bulimia, body dysmorphism, and others which are being revealed as having a strong mental component. Work on the insular cortex and temporo-parietal junction, in particular, has revealed that lesions in those areas disrupt the normal experience and perception of one's own body, and patients who experience body-related disorders like anorexia also

demonstrate abnormal activity in those areas. In other words, while it is necessary to understand the interaction between culture and one's body, it is inconsilient to draw a distinct separation between the mind, body, and culture, and then *ignore the mind* when mapping complex concepts like sexual or racial identity which are certain to be composed by all three components in concert.

An embodied perspective would decisively *not* claim that cultural and social conditions have no effect on the body and the mind; indeed, Raymond Gibbs Jr. explicitly said as much: "Bodies are not culture-free objects, because all aspects of embodied experience are shaped by cultural processes. Theories of human conceptual systems should be inherently cultural in that the cognition that occurs when the body meets the world is inextricably culturally based" (*Embodiment* 13). Moreover, as Edward Slingerland writes, "the recognition that a large part of the environment in which humans find themselves embodied is itself a human creation has focused attention on how *cultural differences in embodied experience affect thought*, as well as how cultural forms are created and transmitted" (13). Cultural interaction is clearly alive and well in the embodied perspective. However, as so many embodied psychologists and critics of post-structural and postmodern socially-based literary criticism have pointed out, the body and the mind are also decisively *not* a *tabula rasa* waiting for social and cultural inscription. For literary theorists interested in the social and cultural effects upon embodied identities, whether sexual, gendered, or racial, the embodied perspective would allow those same theorists to propose their hypotheses within a consilient paradigm where their hypotheses could be scrutinized in light of relevant empirical evidence. This is a crucial component of establishing a consilient cognitive approach to literary studies.

A consilient, embodied approach to literary theory would still be able to explore the "ways in which culture and language shape the human mind, pointing to work from anthropology and cross-cultural psychology that suggests how diverse cultural training, environmental variety, diversity in modes of production and social organization, and the effects of entrenched cultural forms and metaphoric blends can retune or alter the basic universal perceptual and conceptual structures" (Slingerland 23). In other words, consilience with embodied theories would allow literary theory to continue to practice within the well-established sphere of concerns in which it already travels, while simultaneously lending our theories a methodologically-based, interdisciplinary credibility.

Part IV. The Mind in Literature and the Literary Mind

An embodied perspective would not only help make established literary paradigms consilient with the relevant sciences, it would also open new lines of investigation and research within literary studies. While there are any number of areas with an overlap of interest between the cognitive neurosciences and literature, perhaps the most relevant area with the most public interest is emotion. Once banished from the sciences, emotion has returned as one of the most highly researched areas within cognitive neuroscience. It is now being addressed from multiple vantage points and concerns. Autism is now at least partially understood as an emotional disorder. Normal cognition has been re-wed to emotion. Empathy and imitation are being revealed as foundational to the development not only of learning, but of social intelligence, and, as it has been suggested by several mirror neuron researchers, perhaps language itself. Likewise within literary studies, emotion has returned as an important object of study.

Richard Gerrig, Marisa Bortolussi and Peter Dixon, Raymond Gibbs, Jemeljan Hakemulder, Wayne Booth, and Martha Nussbaum have all argued the study of affect, (whether in empathic response to characters, ethical evaluations, or as an emotional response that occurs in reading), has been weeded out of literary studies and needs to be re-instituted. Before we turn to the literary argument for studying emotion, however, we must first pause to develop a consilient understanding of emotion within the cognitive neurosciences.

Emotion, as defined by Antonio Damasio, Bud Craig, Ralph Adolphs and other embodied proponents, is the body's cognitive response to changes in its homeostasis; "feeling an emotion is the experience of such changes [in homeostasis] in juxtaposition to the mental images that initiated the cycle" (Damasio 145). Homeostasis is the property of a system, which can be either open or closed to external influences, to regulate and control its internal environment to maintain a stable, constant condition. For biological organisms like people, homeostatic functions are those that regulate heartbeat, respiration, temperature, pain thresholds, hunger, thirst, and the like. In this view, then, emotions are fundamentally body-based, and they absolutely cannot be understood without reference to the bodily structures with which they interact, be it increased heart rate, tightening of the skin, the release of chemicals, or any other bodily response. Emotions are cognitive responses to evolutionarily salient stimuli, and can be thought of as fast (in neural time) responses to potentially dangerous and critical percepts, still cognitive in nature, but without the characteristic reflective character of other high-order cognition.

Joseph LeDoux's work on auditory perception in rats highlights the cognitive differences between emotion and the feeling of an emotion. There are two pathways in the brain that deal with fear-inducing stimuli, one which travels from the sensory thalamus to the sensory cortex before finally reaching the amygdala, the fear-center in the brain, and a second path, in which signals travel directly from the sensory thalamus to the amgydala at about twice the speed of the first path (six one-thousandths of a second). What is the advantage of having two pathways that eventually reach the same place in the brain? I will quote LeDoux's explanation because it not only answers this thorny question, but also because it draws attention to the distinction between an emotion and feeling an emotion that I am attempting to make here. LeDoux writes:

> Imagine walking in the woods. A crackling sound occurs. It goes straight to the amygdala through the thalamic pathway. The sound also goes from the thalamus to the cortex, which recognizes the sound to be a dry twig that snapped under the weight of your boot, or that of a rattlesnake shaking its tail. But by the time the cortex has figured this out, the amgydala is already starting to defend against the snake. The information received from the thalamus is unfiltered and biased toward evoking responses. The cortex's job is to prevent the inappropriate response rather than to produce the appropriate one. Alternatively, suppose there is a slender curved shape on the path. The curvature and slenderness reach the amygdala from the thalamus, whereas only the cortex distinguishes a coiled up snake from a curved stick. If it is a snake, the amygdala is ahead of the game. From the point of view of survival, it is better to respond to potentially dangerous events as if they were in fact the real thing than to fail to respond. The cost of

treating a stick as a snake is less, in the long run, than the cost of treating a snake
as a stick. (163-165)

Emotions, then, are our unreflective, bodily-initiated, fast-acting cognitive responses to evolutionarily salient stimuli. In this view, our feelings of emotions are the slower (again, in neural time), cognitive evaluations of the fit between our emotional response and the actual stimuli. What makes this so particularly interesting for those studying emotional responses to literature is that we must be struck by how such a complex cognitive response is generated by a stimulus that is artificial (i.e. fictional). Literature, as it exists in its physical form, is *not* evolutionarily salient; it does not threaten the body with danger, promise physical pleasure or satiation of biological drives, nor *directly* engage us socially, and yet, despite all of this, it still produces dramatically complex and powerful emotional responses. In other words, the question we should be asking is *how* are simple printed words on a page capable of generating complex emotional - and felt emotional – responses, and, perhaps even more importantly, *why* does this happen?

As I have already suggested, there are a number of literary and psychologically minded scholars at work on pieces of this question. Psychologist Richard Gerrig has empirically studied what he calls the paradox of emotional response to literature – the evocation of real emotions in response to a fictional stimulus – and how that emotional response guides integration of fictional knowledge with real-world knowledge. Raymond Gibbs Jr., another psychologist whose career has often involved examining literature, has clearly demonstrated through a comprehensive series of studies that emotional responses serve as a sort of guide for reader's making judgments about authorial intention[5]. Literary theorists Wayne Booth and Martha Nussbaum have each done extensive work with

emotion as a tool for ethical and moral evaluation of literature. Jemeljan Hakemulder's book, *The Moral Laboratory*, takes Booth's and Nussbaum's argument a step further, as it backs up this claim with an extensive empirical research program that indeed does seem to suggest that reading literature (as opposed to reading essays, and other non-fictional texts) does indeed seem to produce ethical changes. Clearly, within both literary studies proper and psychological approaches to literary processing, there is a empirical movement to study emotional responses to texts gathering strength.

Antonio Damasio's somatic marker hypothesis provides a similar perspective from neuroscience, acting as an embodied challenge to the exclusion of affect from our interactions with literature. In Damasio's view, the source of feelings is the body, and feelings are inevitably about the body, its state, and its potential future states. Somatic markers are automatic, immediate, emotion-driven evaluations that are not necessarily available to conscious reflection. It is important to realize that "[s]omatic markers do not deliberate for us," but , rather, they "assist the deliberation by highlighting some options . . . and eliminating them rapidly from subsequent consideration" *(Descartes' 174).* In this embodied view, then, emotions and feelings are often our first cognitive interaction with a situation, including, I would argue, our interactions with literature. While our pursuit of disinterested scholarship asks that we table our own emotions when evaluating the data that concerns a particular hypothesis, there is *nothing* that forbids us from making our emotional responses *themselves* a part of that which we study. Any reader of fiction is familiar with the experience of being emotionally involved in a story, whether anxious while reading a suspenseful thriller, fearful while reading a horror story, disconsolate while reading a tragedy, or amused by something clever in the book. Similarly, it is no

secret that readers have emotional reactions to characters and plot events. Emotional responses to literary texts are self-evident. The embodied perspective challenges the view that they are unimportant to our understanding not only of literature itself, but to our reactions to it, and also our eventual comprehension and interpretation of it.

But an embodied approach to literature, situated within a larger consilient cognitive methodology, could open up even more productive, directly relevant lines of research for literary scholars. In their book, *Psychonarratology*, Marisa Bortolussi and Peter Dixon laid out a program whose spirit was captured in the subtitle of the book: "Foundations for the Empirical Study of Literary Response." They argue for a true interdisciplinary study of narrative, spanning the various specialized paradigms in "literary studies, cultural studies, linguistics, discourse processing, cognitive psychology, psycholinguistics, cognitive linguistics, artificial intelligence and . . . ethno-methodology and critical legal studies," unified by a methodology which is virtually identical to that which is currently practiced in the psychological sciences (2). Their approach is similar in spirit and goals to my own, and it is certainly one which would qualify as consilient in the main.

Bortolussi and Dixon conclude their book with a list of "unsolved problems" and "other directions" left for the empirical study of literature; they themselves focus mainly on readers' reactions to and processing of textual features, a la narratology (240-254). Among the areas of research left undeveloped, they name "memory and attention," "reading context," "extratextual information," "individual differences," "literariness," and "genre" (240-254). An embodied perspective which attends to emotional responses to literature could provide a nexus for the study of many of these unexplored areas. For

example, there has been relatively little exploration done on the nature of mnemonic effects within literary texts. Questions arise such as what do people remember when they read literature? Why, and how accurate are those memories? These questions are especially relevant to the instruction of literature, and might also provide clues to the qualities which make a work of literature enduring, or, in other words, contribute to its "literariness".

Within empirical memory studies, there is already a wealth of documentation surrounding the well-established relationship between emotion and memory; "emotion can affect the likelihood that we remember prior experiences, and emotion also can influence the types of details that we remember about past experiences" (Kensinger 243). However, emotion can also distort the accuracy of a memory, and as such a distinction must be made between remembering the content of a particular event (or text) and the remembering of the emotion elicited. These complex relationships have been extensively studied by memory researchers, and as cognitive literary studies begins to approach consilience with the sciences, these are the sorts of studies that are essential to conducting our own research. Asher Koriat's work on monitoring and control processes in metacognition, including directed remembering, are especially relevant to the study of literature. One important distinction Koriat makes is between the quantity-oriented versus accuracy-oriented approach to memory. Quantity-oriented approaches to memory are "traditionally used to tap the amount of studied information that can be recovered . . . reflecting the likelihood that each input item is correctly remembered," whereas accuracy-oriented approaches "evaluate the dependability of the memory – the extent to which remembered information can be trusted to be correct" (*Monitoring* 491). In

Koriat's work on metacognition, he has distinguished between "information-based and experience-based processes," which he claims "shares some features with the old distinction between reason and emotion . . . but differs from it" (*Metacognition* 301). The distinction is simple enough to understand: when asked to judge their own confidence in a memory, people use one of two methods to metacognitively evaluate their own cognitive processes; they rely either upon the memory itself – if someone asks me for my mother's maiden name, my judgment of knowing will rely strictly upon whether or not I can recall her maiden name – or they rely on the subjective feeling of the "fluency with which information is encoded and retrieved" – how easy and sure it *feels* to remember my mother's maiden name (*Metacognition* 298). The implications for memory of literature should be clear. Reader's responses to literature are highly emotional, thereby effecting not only *what* they remember, but also how *accurately* (the quantity-accuracy distinction) they remember it. Unfortunately, we as a discipline have little understanding of what textual features, constructions, genre expectations, background knowledge, and individual differences contribute to this process, and almost no evaluative guidelines for understanding how memory for a text might guide interpretation, and, perhaps most condemning, no well-established rubric for evaluating memory for complex literary texts.

An embodied approach to literary study would also allow literary researchers to ask relevant and timely questions concerning the nature of the physical object of the book and its relationship to textual comprehension, affect, and memory. As the rise of electronic publishing and hand-held electronic readers like Amazon's Kindle or Apple's iPad coincides with a growing environmental concern over deforestation, we should not refrain from exploring the actual, embodied relationships between reader and the physical

form through which the text is accessed. Even for "book" purists, there might be differences between the reader's experience of a large, leather bound folio edition of a work versus a cheap paperback version. The embodied perspective holds that there are not only clear phenomenological and experiential differences, but that these differences themselves are likely to influence cognitive, mnemonic, and affective responses to the text. Additionally, as the nature of the "book" itself is changing, it is no less important to note the evolution of the reading context, and what this may mean to the literary experience. Again, an embodied perspective would posit vast cognitive and experiential differences between reading in a quiet, fire-lit den while seated in a comfortable armchair versus a well-lit office reading from a computer screen or reading laying on one's stomach on a futon in a dark, noisy dorm room or even in the controlled, slightly unnatural and uncomfortable environment of the classroom. What those exact differences are and if they are negligible or not are empirical matters, and despite the direct relevance and pressing nature of these concerns, have been mostly overlooked by literary scholars.

By this point, it should be clear not only what a embodied approach to literary studies means, but also why it is important for literary scholars who normally think of their object of study as divorced from the body and beyond a scientific approach. However, within literary studies there is already a growing movement of scholars working within a cognitive and/or embodied approach. This is an encouraging trend, yet, in order for the cognitive movement (or the embodied movement, or the evolutionary movement) to become more than just another place holder in the temporal succession of interpretative paradigms within literary studies, the movement, as a whole, must be made consilient with the relevant sciences from which it is drawing.

Part V. Making the Cognitive Revolution Consilient

Similar to my focus on the embodied movement within contemporary psychology, my focus on cognitive literary studies will likewise be concerned with cognitive literary theory circa 1985 and after. While there are many works of value produced long before what Alan Richardson and Ellen Spolsky dubbed the "second generation in the cognitive study of literature," it is precisely research from that period on with which I am primarily concerned, for two simple reasons (x). First, I will focus on cognitive literary theory produced starting in the mid-1980's because it is both the easiest to test for consilience with the contemporary sciences. The brain sciences have been dramatically revolutionized since the 1980's. Since I am claiming that cognitive literary theory *must* be consilient with the relevant scientific disciplines, it is easiest to make a side-by-side comparison with the literary theory that has been produced under the current neuroscientific and psychological paradigms. Examination of texts produced prior to this period would call for either a historical approach to their consilience with the scientific knowledge at the time of their production, or updating and cross-referencing with contemporary knowledge, processes that, while valuable, require a great deal of historical research to execute fairly. Second, because cognitive literary theory produced in the last twenty five years should be held consilient with knowledge from the brain sciences produced in that time period, it also provides the clearest examples of the promise of consilient cognitive literary theory, as well as the perils of theory that is produced without the evaluative guidelines of consilience.

It is clearly beyond the scope of this project to evaluate every single piece of cognitive literary criticism produced in the last twenty five years for its consilience.

Instead, I intend to demonstrate the *process* of evaluating a piece of cognitive literary criticism for its consilience, while extending the argument for why consilience is crucial to our methodology. Thus, I hope to introduce a practice for holding literary scholarship responsible for its theoretical consilience and methodical practices so that other, like-minded literary scholars can extend the project of founding a scientific base for the study of literature. The disciplinary changes I am calling for are indeed dramatic, and cannot be expected to happen overnight, particularly as they will have to overcome internal resistance. However, with enough dedicated scholars who wish to employ the methodical practices of the sciences and constrain their theories within the bounds of consilience, the goal of establishing a consilient, scientific literary studies can be reached.

Narrative Theory and the Cognitive Sciences, an edited collection by David Herman published in 2003, serves as an excellent example of both consilient cognitive literary criticism as well as criticism which is "interdisciplinary" only in the sense of its importation, misappropriation, and misapplication of "gratuitous jargonizing" and in its "cherry-picking of only supportive evidence" (Gottschall 85). As such, it is a "good news, bad news" book. The collection serves to illustrate the diverse and vibrant areas of research within the approach, a matter that is at once exhilarating even as it demonstrates the excesses of faux-interdisciplinarity. As Alan Richardson and Francis Steen may have unconsciously suggested in their introduction to a special issue of the journal of *Poetics Today* dedicated to "Literature and the Cognitive Revolution," while the work of cognitive literary theorists and critics "finds its inspiration, its methodology, and its guiding paradigms through a dialogue with one or more fields within cognitive science,"

it sometimes holds little more in common with those same scientific fields than its inspiration (1).

Herman himself introduces the collection, and traces the development of the "cross-disciplinary narrative turn" to a cognitive paradigm to three major research initiatives: "the rise of structuralist theories of narrative in France . . . the advent of the sociolinguistic study of personal-experience narratives . . . and the focus on narrative by cognitive psychologists and artificial intelligence researchers concerned with story grammars and with scripts and schemata" (5). While these three movements were indeed foundational to the interdisciplinary study of narrative and literature, they are all movements that occurred *before* the rise of the embodied movement in contemporary psychology. Therefore, some of their theoretical assumptions, particularly those which hold that the mind is a computational device, to be studied only in terms of information processing without need to reference its substrate (biological or digital, etc.), or the environmental context of thought, have been challenged by the development of embodied cognition. The first lesson to be learned is: **Consilience is not a static state, but a dynamic process**.

In the seven years since the publication of Herman's collection, the embodied movement has gained significant traction within psychology and evolutionary biology. The challenges it has raised against computational models of the mind are now being *integrated* into existing research paradigms. For cognitive literary theorists, then, consilience must consist of a continuous updating of theories and ideas, to reflect the changes and developments in the fields that originally inspired those theories. Only by making a concerted effort to continually revise our ideas in light of new evidence in favor

of or in conflict with those same theories can cognitive or any other branch of literary criticism become consilient. This progressive quality of knowledge is something with which scientific practioners are already conversant. New hypotheses are generated in the light of knowledge produced by past experiments, and old hypotheses and theoretical structures are discarded as necessary.

This is a practice that literary studies needs to emulate. For example, contemporary literary theorists still construct theoretical models of the mind or cognition based upon the work of Sigmund Freud, a theorist whose ideas have long since been discarded in the psychological sciences. Likewise, the practice of having to engage current literary theories and hypotheses with past paradigms, simply out of a misplaced anxiety of influence, is one which does not benefit the progression of knowledge within the field, nor make our scholarship more accessible to the uninitiated. If a consilient minded literary scholar wishes to explore the gendered nature of literature, he or she should not *have* to reference the vast amount of gender-focused literary criticism produced in the past thirty to forty years, particularly the criticism that is itself inconsilient and methodologically unsound. That body of work, while *potentially* valuable for its insights, is empirically groundless, and the only reason for a consilient scholar to return to it is to begin to test the vast number of truth claims made without empirical support. Consilient literary theory would replace this anxiety of influence with a structured development of data, hypotheses, and theories, allowing for a more progressive interaction within the field.

To return to Herman's collection, I should note that it is presented in four sections, and each of those sections marks off a major approach within the cognitive

literary paradigm: Approaches to Narrative and Cognition; Narrative as Cognitive Endowment; New Directions for Cognitive Narratology; and Fictional Minds. Approaches to Narrative Cognition and Narrative as Cognitive Endowment are essentially looking at the same aspect of human cognition as it relates to narrative from opposite ends of the spectrum. Psychologists and computer scientists like Richard Gerrig, Giovanna Egidi, Kitty Klein, William Frawley, John Murray, and Raoul Smith focus on the effects that literature and narrative produce upon readers and the textual features and techniques of construction that give rise to those effects. Cognitive linguists like Mark Turner, and cognitive literary theorists like H. Porter Abbott and David Herman, on the other hand, theoretically posit how narrative itself is central to and structures all of human cognition. In other words, the scientists empirically study and quantify how readers are effected by reading narratives, while the literary theorists theoretically claim how narrative structures readers' thoughts and thought processes even when not reading. It is striking to note the different methodologies employed in the two sections; each of the papers in the first section is empirically supported by experiments and observations conducted by the scientists, as well as prior research of other scientists; each of the papers in the second section is argued entirely discursively, with references to other theoreticians, but no move is ever made to actually support or demonstrate the structures at work that they are claiming exist at a fundamental human level. The second lesson to be learned is: **If science *is* the method, beware of those arguments where the method is missing.**

I want to be clear that this is not a blanket statement equivalent to something like the following: if its not empirical, quantifiable, observable, or measurable its no good to

study it. Nor am I trying to say that there is no room for or value in strictly theoretical hypotheses. What I will go on record as saying is that if purely theoretical hypotheses are put forward, they should be tested. The problem for literary scholars is then two fold. One, and perhaps the least troublesome, is that there is a disciplinary practice of simply referring to the master argument as evidence for one's own argument. Another person's argument, no matter how well made, should never qualify as evidence. What this does, as has been well documented, is establish a tradition of authoritarianism in which what the master says is true, a la Freud, Lacan, Marx, Foucault, Derrida, and so on. Evidence for a hypothesis *is* empirical, observable, quantifiable, and/or measurable. Which brings me to the more problematic issue facing literary scholars. Because we have no disciplinary tradition of studying literature in this way, and because the nature of the object which we study is so complex, it is going to be difficult to create sensible rubrics and tools for our evaluations. How *does* one measure "literariness," and *with what*, and, better yet, *why don't we ask these questions*? A great deal of what we do in the early stages of making literary studies consilient is going to consist of borrowing from the more established sciences, which is fine in the early going. What we should not shy away from, however, is the development of our own measures and tools, particularly once our investigations have begun to present a more nuanced understanding of what the study of literature actually entails.

To return to *Narrative Theory and the Cognitive Sciences*, the third and fourth sections of the book are intended to show the application of the cognitive paradigm to particular works of literature, again from two directions. Part three takes works of literature as specific instances of a broader category, and then works from those instances

to try and develop any number of common narrative elements, structures and properties: Marie-Laure Ryan's empirical study of the development of readers' spatial representations in Gabriel Garcia Marquez's *Chronicle of a Death Foretold* is a prime example of this kind of work which focuses on specific readers engaging with a specific text which reveals cognitive processes that can be generalized upon further testing and refinement. The fourth section turns to literature for representations of the cognitive structures and processes that define consciousness as it is understood by the contemporary brain sciences. As we saw with the book's first two sections, there are distinct differences in the methodological and theoretical approaches taken by different authors that will allow for the further demonstration of what consilient cognitive literary theory must, can, and cannot do.

Manfred Jahn's article "'Awake! Open your eyes!' The Cognitive Logic of External and Internal Stories" in section four of Herman's book is an example of methodological failure. Jahn's argument proceeds by making use of three literary examples – Billy Wilder's *The Apartment*, Coleridge's "Kubla Khan," and Richard Wagner's opera, *Ring* – to demonstrate both the importance of internal stories to external stories, as well as the very *existence* of internal stories. Jahn defines an external story as "stories which exist in some physically tangible form," and internal stories as those stories "which are stored in memory and performed in the mental theater of recollection, imagination, and dream" (195). The problem is with Jahn's methodology. Like many literary theorists, Jahn takes literary examples as *evidence* to support his hypothetical model for the generation and interaction of internal and external stories in *actual human cognition*. This is an error made by far too many cognitive literary theorists and one

which demands special attention. Literary examples are not evidence of *anything* that occurs in the minds of real people, *excepting* their status as a specific type of cognitive artifact that human beings produce and cognitively interact with. Cognitive literary theorists cannot build hypothetical models of actual human cognition and then turn to literature for their evidence. They should, instead, develop and carry out experiments with real readers in order to gather empirical observations in support, refinement, or refutation of their hypotheses. As the title hinted, Jahn's article is entirely premised on logical grounds, and while intuitive, fails at being methodologically sound.

Marie-Laure Ryan's "Cognitive Maps and the Construction of Narrative Space," on the other hand, is exemplary both for its methodological and theoretical consilience. In stark contrast to Jahn whose theoretical formulations (the internal and external story) are novel and unconnected, Ryan develops a previously established concept from psychology: the cognitive map. Ryan stages her argument as a development of the psychological and philosophical idea of the cognitive map, before providing specific questions her research seeks to explore. Jahn offers up his theory as a whole, a puzzle already solved, while Ryan, more in the scientific spirit, offers a set of hypotheses that her essay will attempt to answer:

> "What are the relations between cognitive maps and graphic maps? To what extent and in what detail do mental maps of textual worlds need to represent spatial relations between objects? Through what strategies do texts facilitate the conceptualization of these relations? Is a totalizing, bird's-eye-view mental image of narrative space necessary to a proper understanding of plot, or do readers work from cartographic fragments?" (216).

The formation of specific hypotheses that can then be empirically supported or refuted is a crucial step for consilient cognitive literary studies to take. The distinction between Jahn's methodological process and Ryan's is clear. Jahn starts with "evidence," his literary examples, that he then works inductively from to generate his theoretical concepts, the internal and external story. Inductive reasoning is not the problem here, nor is the use of a literary example as a *starting* point to develop theories about actual human cognition. What Jahn fails to do is then develop testable hypotheses that could empirically verify his theories in actual people. Ryan, by starting with testable hypotheses, avoids the error of presenting her initial observations as "evidence," and, instead, moves from her hypotheses to the development of an experiment to test them.

Literary scholars need to receive training in the design, development, execution, and manipulation of experiments in order to become consilient methodologically. Ryan's essay is an excellent example of what controlled experiments can reveal to literary scholars, and how the field as a whole can benefit from the knowledge they produce, as well as the very practice of experimentation itself. Ryan's experiment is a simple one: after developing a detailed "master map" of Gabriel Garcia Marquez's *Chronicle of a Death Foretold* that represented plot actions, object and place relations, and character movement, Ryan turned to sixty high school seniors who had recently finished reading the novel in class and asked them to "draw pictures of the topographical layout" of the novel (224). Ryan stresses that this was not a strictly scientific experiment – there is in fact no manipulated variable, and no controlled condition – but, rather, "an informal attempt to probe into readers' memory and imagination" (224). From the sixty maps that she collected from her subjects, Ryan then looked at patterns in what was represented,

what wasn't, and the types of representation that seemed most important, before finishing her essay with a conclusion that returned to original research questions, drawing upon the evidence she herself collected.

Ryan is correct in her assessment of her study; it is, in fact, not a proper scientific experiment, and there are several serious methodological flaws. However, the data she generates, and the general process she follows is far closer to a scientific process than that which Jahn employs. Ryan has no prior-established criteria for the evaluation of the maps, provides no statistical analysis of the trends she then discusses, and, as previously stated, uses no manipulated variables, nor controlled conditions. Nonetheless, she gathers valuable information which will allow further studies to build upon her data. I have repeatedly stressed that one of the benefits of a consilient research program within literary studies is that it would help develop progressive knowledge that built upon previous literary studies. This progressive feature of consilience is what is at stake with properly executed experiments and data collection, and studies like Ryan's are an initial step in the right direction.

There is nothing a consilient literary scholar can do to affirm, develop, or refute Jahn's ideas, except start fresh by asking, "Do internal and external stories exist in normal human cognition?" and then set about developing hypotheses which construct and predict how those concepts would manifest, when, under what conditions, etc., before finally devising an experiment to verify their existence. In other words, aside from providing a concept that another researcher might choose to actually verify, Jahn's essay provides little in the way of what could be called progressive knowledge. Ryan's essay, even with its methodological flaws, presents a clear series of hypotheses, an experiment that could

easily be replicated, improved upon, controlled, and manipulated, and findings that are presented based upon the evidence she collected that, if verified, could form the basis for a larger research program into the nature of readers' constructions of narrative space, or, if partially or entirely refuted, could still have served as the platform for the research which disproved it. In either case, Ryan's article provides concrete knowledge that can be developed and tested, knowledge that is progressive *even if incorrect*. This is the true beauty of scientifically conducted research: it is impossible for it to be meaningless. Data is data, and while hypotheses, interpretations, and conclusions can be incorrect, data isn't subject to the same problems. It may take a great deal of time to assemble a complete picture of what happens when we read literature, but the *only* way it can happen is through consilient research, conducted through scientific experiments which gather observable data that can then be developed into theoretical frameworks which themselves progressively accumulate, and are refined over time.

Part VI. The Big Picture

What this examination of David Herman's *Narrative Theory and the Cognitive Sciences* has hopefully demonstrated is that while there are any number of ways to approach the study of literature, in order for cognitive literary criticism to make a lasting contribution (to both literary studies as well as the larger scientific community), strides need to be made to ensure both theoretical and methodological consilience. Taken as a whole, the collection serves as a facile representation of the four possible positions of study within cognitive literary studies. The analysis of each of those sections can then serve as a template for addressing specific works of cognitive literary criticism. If cognitive literary studies may be generally thought of as being composed of four

positions made up of two poles each, with the two main objects of study, the text(s) and the reader(s) composing one pole, and the direction of effect between the objects of study composing the other, then we may begin to address larger, more thematic concerns for consilient cognitive literary theory. Figure 1 is a simple representation of each of these poles.

Figure 1

```
        (1) ------------------------------------------------> (2)
      Reader(s)                                              Text(s)
        (4) <------------------------------------------------ (3)
```

Studies in the first two categories would be interested in how the cognitive, cultural, demographic particulars of a specific reader effects textual reception and interpretation. Studies in the first category would focus on the reader, and what the reader brings to the text prior to reading it, and how those differences can be used to explain certain trends within established interpretative communities. Examples of studies at this pole could be research on background knowledge that is brought to bear on the text prior to reading (see Chapter 4 of this project for an example of just such an experiment), linguistic norms that influence reading of texts (such as dominant metaphors which can vary through culture, subculture, historical era, or geography), denotative and connotative usage of particular words, and literary or other training in critical reading skills. Studies conducted from the second position would focus on specific textual features that, once read, evoke responses in individual readers because of who the reader is in terms of demographic make-up. For example, researchers interested in the portrayal of race in Southern fiction could examine how specific characters, plot developments, scenes, and

writing styles interact with different readers according to the reader's background. A hypothesis along these lines could attempt to answer whether race, gender, age, sexuality, life experiences, etc., matter when brought into contact with specific textual features of a certain kind. The major portion of existing socially-oriented literary criticism is essentially interested in this particular area, and needs only to be subjected to methodological and theoretical consilience to be made relevant.

Studies in the second two categories would, in contrast, focus on effects initially propagated by the text and then transmitted to the reader. In position three, research would focus on generally occurring textual features, such as grammatical or lexical styles, genre conventions, and narrative forms, and what these structures can mean to a reader before they've even read the text. Research conducted from position four would then focus on the effects that manipulation of these textual features produces within specific readers, such as cognitive and emotional responses, mnemonics for textual features, and moments that invite the reader to engage in interpretation. These sketches are by no means complete, nor are they meant to be. They are intended to suggest the differences between certain *types* of literary studies in order to draw attention to the common demands for methodological and theoretical consilience that exist regardless of which position a researcher occupies.

Research that focuses on the reader, a line of investigation that is often neglected within literary studies, can follow one of two major directions as indicated by the diagram. If it follows the flow of effects moving from text to reader, then these studies should primarily be concerned with identifying features and structures of narrative and literature that produce certain types of effects on *individual readers*, effects that are

statistically reliable in their occurrence, which can then be generalized to form a picture of a what happens at an aggregate level when a reader addresses a particular scene, structure, style, genre, etc. Richard Gerrig, for example, has studied the textual features employed in detective fiction and suspenseful novels that induce what he calls "anomalous suspense": "to experience suspense with respect to an outcome about which [the reader] should not have any uncertainty" (158). Gerrig studied readers in situations where they knew the outcome of the story they were reading (or watching), most often experienced when encountering a story which has been previously read or watched, or in experiencing a historical narrative where the outcome of the story is a matter of established historical fact.

While his findings should be of great interest to scholars of literary studies, what is more important to this project is his methodology. As a psychologist, Gerrig used the methods of science; he observed individual readers and their reactions, made use of their reported phenomenological experiences, and, perhaps most crucially, *he manipulated the texts and the situations which he was studying*. As Marisa Bortolussi and Peter Dixon have argued "the best technique for assessing such causation [of reader constructions caused by particular features of the text] is to conduct *textual experiments*, in which particular features of a text are identified and manipulated by the researcher" (51). Observational data alone can not prove causation, which is the ultimate goal of scientific study; it can only demonstrate correlations between variables. For causation to be demonstrated, the related variables of a given situation (experiment) must be selectively isolated, and then changed (manipulated) in specific and controlled ways that can be used to predictably alter the results of that situation.

For literary studies, this is a particularly tricky point. The variables being studied are inevitably the following: 1) the text; 2) the reader; 3) the reader's responses to the text; 4) the reading task; 5) the reading environment. Traditionally, literary studies has focused solely upon the first three variables, mostly ignoring the effects of *how* and *why* (task)*,* and *where* (environment) someone is reading has upon their response to the text. This oversight is important to correct, not only because the reading task and reading environment are certain to interact with the reader and the text, but also because they can be easily manipulated within experimental settings. In the detailed studies of the interactions between text, reader, and reader responses that are the bulwark of literary studies, there is little manipulation of those variables. In order to understand how a text actually causes certain responses in a reader, those features of the text thought likely to be responsible need to be selectively identified, and then altered in order to see if the effects produced are indeed dependent upon those features. The difficulty with this approach is that the study of literary texts is the study of texts of such complexity that it is not only difficult to selectively isolate textual features at a reductive level, but also that the alteration of a word, sentence, or even paragraph may not matter much in relation to the whole. Nor is the detailed, empirical study of an entire novel going to be experimentally realistic. This is not, however, a license to stop trying, only a sympathetic challenge to literary studies to develop a method for selectively manipulating literary texts in a way that is methodologically sound and operationally feasible.

The reader, on the other hand, provides a wealth of variables that can be selectively manipulated in order to study their effects on textual interpretation. Classic literary concerns like gender, race, sexuality, ethnicity, religion, etc., can be isolated prior

to reading, and reader responses to the same texts and textual features can be contrasted across those lines. Beyond that, as I will investigate later in this work, background knowledge of the text itself, its author, knowledge of aspects of socio-political history relevant to the text,genre expectations, literary conventions, and period styles, can all be selectively manipulated in order to demonstrate their effect upon readers' responses to a text. The take-home message here should be clear: the selective manipulation and control of variables is absolutely critical to identifying the features and processes at work within the literary experience. Anything less than this practice is scientific-posing without the actual science. Work within discourse processing and other psychological approaches to textual response is already well grounded in its methodology, as it is primarily conducted by people who are trained in the methods of empirical psychology. Literary scholars who are familiar with the characteristics of genre, styles that were prevalent during certain literary periods, and formal structures of literary narratives are well placed to enter this field. They can easily contribute to the cataloging of features and recognition of particular effects brought about by those same features, and begin to construct experiments that manipulate those variables.

The second direction of effect in research which primarily concerns the reader, from reader to text, is in a rather different state of development. For those theorists who are concerned with the reader in terms of what the reader cognitively brings to a text *before* he reads it (i.e. how does narrative and metaphor structure our thought in general) there is the difficulty of empirically proving the claims which are made. For example, the work of cognitive linguists like George Lakoff and Mark Johnson, and Mark Turner and Gilles Faulconnier, are well developed theoretical edifices that are concerned with how

metaphoric and narratological structures underly human language and thought. While both teams of researchers have plenty of naturally occurring examples to demonstrate what they mean, the difficulty here is *how* this area is currently being researched. As part of the quadratic division that is cognitive literary studies, this is the area that is closest in its object of study to the cognitive neurosciences, and, as such, needs to follow those methods if it seeks to *prove* its claims. If these processes are in fact central to human thought than it should be possible to study them in operation through well designed electrophysiological studies, structural and functional imaging, or through individual case and group studies of lesioned and non-normal patients suffering from deficits with these kinds of processes as well as by inducing "virtual lesions" with transcranial magnetic stimulation. While the theoretical edifices built by cognitive linguists and other researchers within this area are enticingly interesting as they relate to our practices of reading, writing, and interpreting literature, and while they are *seemingly* intuitive and well-grounded in observations, until these theories have been connected with the body of work from the cognitive neurosciences which would support, refine, or refute them, they will be in danger of being inconsilient humanistic theories masquerading as scientific hypotheses.

For literary scholars who are interested in the more traditional object of literary studies, the text, there are likewise two directions of effect, and two separate categories of considerations. For those interested in studying a particular text for what effects it produces on the aggregate readers, focus needs to be placed upon specific textual features that are unique to that text, and large numbers of actual readers need to be sampled for the relevant reactions to those features. This is perhaps where traditional literary

interpretations most naturally fit. If one considers the accumulation of interpretations for a single work, single author, period of literature, genre, or any other category within traditional literary studies, what we already have is a large amount of data demonstrating how *trained* readers react to specific features in specific texts, and, perhaps most interestingly, how those same reactions change over time. What we don't have are hypotheses concerning *how* those interpretations are generated, or by what complex matrix of influences, which would have to include training, socio-historical context, individual differences in demographic make-up as well as background knowledge, among a vast array of possible other factors.

That the complexity of individual responses to literature is staggering is self-evident. However, every major scientific advance in knowledge has begun as a question of overwhelming and seemingly insurmountable complexity. A consilient approach to literary studies could make valuable use of the amassed data held within literary interpretations to produce hypotheses that correlate interpretations with interpretative communities. However, this is really only half the story. As we do not have the control group's readings of those same novels – reactions from untrained, non-literary readers – it is difficult if not impossible to know to what degree the reading processes and interpretative practices of trained literary critics can be then generalized to the normal population. While the production of more literary interpretations is certain to continue, and as some scholars have suggested at a pace that is already beyond critical mass, what we need to do is to begin creating procedures for evaluating the massing database of literary interpretations, and then develop a framework for comaring how regular readers react to those same texts over time.

Literary studies is also replete with studies focused on the text which partially detail the effects which move from text to reader. Once again, the studies amassed about genre conventions, narrative forms, period styles, poetic and prosaic grammatical and lexical constructions provide a wealth of data that could be used to start asking the more relevant questions about *how* these same conventions, forms, styles, and constructions actually produce certain cognitive effects. To return to Richard Gerrig's study on anomalous suspense as an example, while Gerrig has demonstrated the phenomena of anomalous suspense as existing for individual readers, and thus an aggregate "reader," what remains to be done to completely formulate this *one single response* is develop the connections between specific genres (like detective fiction, mystery novels, thrillers, and historical fiction) develop the conventions used in those genres, and ask how those genres and conventions, or manipulations of those genres and conventions, actively contribute to producing anomalous suspense in readers.

While I have mapped out the general territory that consilient cognitive literary studies can be expected to cover, I have yet to specifically address the research that has been done and is being done within contemporary cognitive literary studies, and how it does or doesn't fit in with the program I've outlined here. Despite being a young approach, as well as being primarily conducted by a relatively small number of researchers and theorists, the range of those works is already itself the topic of a book-length project. It is not my intention to reduplicate that information here. Instead, in order to best demonstrate the continued call for methodological and theoretical consilience, I have chosen what I consider three well-known examples from within the field to

represent the different methodological practices that are (or aren't) followed, and the results that those practices have upon the level of consilience of the scholarship.

In 2004, Alan Palmer published *Fictional Minds*, a book that Palmer himself opens by claiming it is about "the amount, range, variety, and reliability of the information on the fictional minds of people in books that we are able to obtain from those books" (1). The Introduction to Palmer's book promises to "illuminate the study of fictional minds by making use of the insights of some of the disciplines relating to real minds," the very claim that I have also made here, and will continue to make throughout this project (4). Most of Palmer's book is an attempt to synthesize various positions and discourses on the contemporary understanding of the mind. After using two chapters to describe the classical narratological approaches to the study of fictional representations of consciousness, as well as a discussion about the speech act and speech category debate, Palmer moves on to summarize the "whole mind" as understood by "the *parallel discourses* on real minds, such as cognitive science, psycholinguistics, psychology, and the philosophy of mind," the social mind, and the fictional mind (14). Where things take a dramatic turn for the worse is in chapter seven, "The Fictional Mind in Action," where he maps out his own theories as developed from his understanding of the synthesis presented in the proceeding chapters.

To begin with, the most notable omission within the chapter (and thus, the whole book) is any effort to empirically support *any* of his claims. Instead, after *saying* how important the "parallel discourses" of mind are to understanding literature, he simply returns to conventional literary close reading. Palmer's chapter sub-headings reveal this bias towards speech act and narratological concerns: "Inner speech, direct thought, and

free indirect thought" and "Contextual thought report." The little that would qualify as original hypothesizing occurs in the sub-headed sections entitled "Intermental Thought" and "Doubly Embedded Narratives," but again, Palmer fails to provide any evidence for their existence beyond the one or two literary examples he gives, nor does he ever address their relationship to the cognitive discourses he has previously claimed to be so important to understanding these concepts. Instead, he makes the move all too many literary theorists make, which is to appeal to the work of other literary theorists and philosophers, and to conveniently forget the challenge of the scientific method and the knowledge base of the relevant sciences when it no longer supports his hypotheses. In the "Intermental Thought" section he turns to "the possible-worlds paradigm," a philosophical paradigm that allows him to explore the possibilities of "collective consciousness"; it is no coincidence that after citing his two literary "experts" in possible-worlds theory, he cites science-fiction novelist Phillip K. Dick as a further authority on the phenomena of "group mind" (219).

This is precisely the sort of practice that plagues literary studies. In his own book, Palmer has argued for the importance of interdisciplinary understanding, and within that same book, he has turned away from everything that those disciplines *do* and *know* in order to explore what is a more provocative, albeit quasi-fictional, theory. While the notion of a "collective consciousness" or "group mind" *may* turn out eventually to have some scientific basis, as we are still so far from understanding the relatively "simple" dynamics, structure, and processes of the individual human brain when it is engaged in the most mundane and fundamental of cognitive tasks (like simple visual perception), it seems ridiculous to jump from the sciences which attempt to develop a clear

understanding of the mind in those processes, to what is almost certain to be an ill-formed question: What exactly is a "group mind" or "collective unconscious" and how could one go about studying it in action?

When I have written about inconsilient practices, or others have warned against sampling scientific jargon and data for confirming evidence and interesting theories to make use of, this is the sort of practice that is being warned against. This is *not* acceptable as a consilient practice. Truth claims, such as Palmer makes, demand evidence, and support from other literary theorists who themselves have no evidence of what they claim, or from literary narratives which are, after all, fictional, are not sources of evidence. While I *do* think that exemplification in literature is important for the development of a nuanced and lived understanding of scientific concepts, literary exemplification is *not* a source of evidence for our hypotheses. In order for literary studies to distance itself from this argumentative strategy, marked by unscientific methodology, Palmer's sort of "scholarship" should be recognized as unacceptable.

In 2006, Lisa Zunshine published *Why We Read Fiction: Theory of Mind and the Novel*, a book that was a large factor in her being awarded a prestigious Guggenheim fellowship the next year. The book, as its title suggests, draws its primary conceptual organization from the work of Simon Baron-Cohen, a neuroscientist most noted for his work with autistic patients. "Theory of mind" is, itself, an incredibly simple concept; it refers to "the ability to represent the mental states of others, i.e. their thoughts, desires, beliefs and intentions" (Ward 325). Proponents of "theory of mind" as a cognitive process often focus on what is called "false belief," which is a "belief that differs from one's own belief and that differs from the true state of the world" (Ward 325). To test theory of

mind, Baron-Cohen developed the Sally-Anne task, a task used in diagnosing autism in children via the ability to hold false beliefs. In the task, the child interacts with two people, Sally and Anne. Sally then puts a marble in a basket so that Anne can see what she has done, before Anne then leaves the room. While Anne is away, Sally then moves the marble from the basket to a box nearby. When Anne re-enters the room, the child is asked, "Where will Anne look for the marble?" Normal children over four years of age and retarded children alike will answer "in the basket," whereas children with autism will answer "in the box." It is important to understand that this is not a failure of memory, but of attribution of belief. When asked, autistic children are as good as normal controls at remembering the marble's initial position. In Zunshine's own words, her hypothesis "suggest[s] that fiction engages, teases, and pushes to its limits our mind-reading capacity" (4).

Where Zunshine succeeds is in taking contemporary hypotheses within the relevant cognitive and psychological discourses and making deft and appropriate use of them in explaining cognitive responses to fiction. After laying the ground work for her conceptual analysis through the work of Baron-Cohen and likeminded "theory of mind" proponents, she uses the work of Leda Cosmides and John Tooby on metarepresentation and source monitoring to describe how fictional narratives "rely on, manipulate, and titillate our tendency to keep track of *who* thought, wanted, and felt what, and *when*" (5; original emphasis). She ends her book with an examination of detective fiction and the "recurrent features of this genre" to "explore the exaggerated literary engagement with our source-monitoring capacity" (5). My purpose here, as it was with Palmer's book, is not to provide a detailed review of her book, nor critique or applaud her for her choice of

conceptual focus, but to engage her project with an eye focused on its consilience. In regards to its depiction of the theoretical ideas of "theory of mind," Zunshine's book is exceptionally accurate. She takes little liberty with the concepts as they exist within psychology when she applies them to literature. For any reader who is interested in seeing the power of what cognitive literary studies can reveal about how and why we indeed might read fiction, I do recommend Zunshine's book. However, in a consilient cognitive world, all is not well.

Where Zunshine fails is that she never once clearly formulates a testable hypothesis, despite all but explicitly suggesting one, nor does she ever bother to collect data in any other way than the traditional method of close reading, which is unfortunate, considering that her analysis of detective fiction all but suggests an empirical study. She also commits the sin of cherry-picking her supporting evidence. While "theory of mind" has a number of influential supporters within the cognitive neurosciences, most notably Chris and Uta Frith, and Baron-Cohen, it is only *one* of two competing explanations for how the human brain achieves the representation of other people's minds. The other is simulation theory, primarily put forward by the eminent Italian neuroscientists who together discovered mirror neurons: Giacomo Rizzolatti, Marco Iacoboni, Vittorio Gallese, and the philosopher Alvin Goldman. Similar to her selection of confirming evidence for her theories about fiction in relation to theory of mind, Zunshine also overlooks a large body of work on metarepresentation and source monitoring within empirical psychology, particularly within memory studies. Zunshine's shortcomings are particularly frustrating because, unlike Palmer, she represents a faithful use of scientific knowledge within cognitive literary studies, but does so without ever really attempting to

change her methodology. What is perhaps even more frustrating is that the few remaining steps needed to make the book actively consilient in its approach are all but made within the book.

Zunshine all but explicitly states a testable hypothesis at several different points in her book. For example, in the section in which she examines detective fiction's engagement of our metarepresentational abilities, she claims that "detective stories 'work out' in a particularly focused fashion our ability to store representations under advisement and to reevaluate their truth-value once more information comes in" (124). This is, in fact, a truth claim, and all it would need to become a testable hypothesis is to suggest a way of testing this truth claim. If, as Zunshine suggests, detective fiction does indeed strengthen our mental metarepresentational faculty, then she should suggest a method for measuring that faculty, and then measuring the effects of reading detective fiction versus other controlled types of texts on its development. Her argument is intuitive, but as it is, completely unsupported by any empirical evidence. The most important step in establishing a consilient cognitive literary studies is to radically alter our methodology to become more quantitative and empirically based. In Zunshine's case, it is troubling to see a scholar who takes as great pains as she does to remain faithful to the scientific concepts with which she is working, take no additional step to then test and verify her own theories[6]. This is the cardinal sin of literary studies. Cognitive literary studies needs to take conscious strides to alter its methodology, or it will be just another interpretative paradigm which produces truth claims with no testing or confirming evidence to back them up.

The good news, however, is that there are a growing number of researchers within literary studies (and not just cognitive literary studies) who are taking great pains to ensure that their work is both theoretically and methodologically consilient. Franco Moretti, Jonathan Gottschall and Joseph Carroll lead a number of evolutionarily minded literary theorists whose work is methodologically consilient, as it is empirical and quantitative, as well as theoretically consilient with work being done in contemporary biological and social evolution studies. Within cognitive literary studies, if we exclude the work of Richard Gerrig, Art Graesser and other psychologists who study discourse processing often with an eye towards literary reading, the work of David Miall, Jemeljan Hakemulder, Marisa Bortolussi and Peter Dixon are all shining examples of both theoretical and methodological consilience. Bortolussi's and Dixon's book, *Psychonarratology*, provides a stark contrast to Palmer's book in terms of their consilience. Like Palmer, Bortolussi and Dixon extensively cover the historical developments within literary theory that have led to the current cognitive approach, including, again like Palmer, sections on narratology and speech act theory. What separates them from Palmer, and makes the project consilient, is that they then 1) provide theoretical support taken from the related scientific disciplines, most notably discourse processing and psychology 2) carry out methodological consilience by both identifying textual features which give rise to the effects they are proposing, as well as suggesting (hypothesizing) how those features generate said effects, and taking the critical step of *providing the empirical support for the existence of those effects*, either through studies carried out by other researchers, or, much to their credit, by relating the evidence found through their own experiments and empirical studies.

While readers accustomed to the dramatic, sweeping theoretical claims that literary theorists like Palmer often make within their work may find Bortolussi's and Dixon's findings intuitive and perhaps even simple, what it is important to note at this point is that Bortolussi and Dixon *have found* something, and, in doing so, have provided a step for researchers who wish to advance the study of narratological features within literature to do so. Even if their findings turn out to be incomplete, misinterpreted, or even blatantly incorrect, because of their theoretical and methodological consilience, their studies will still be of direct value to the production of scholarship which seeks to extend, refine, or replace their work. Alan Palmer's work, because of its inconsilience with any theoretical developments within the psychological or social sciences, and because his methodology leaves his theories entirely unsupported, is almost certainly relegated to the fate of most literary scholarship: obsolescence.

What is hopefully clear by this point is how much work has yet to be done within consilient literary studies. However, I have yet to really address how that work is relevant to the world beyond the university halls. It is an accepted truth that the study of literature is itself a worthwhile endeavor. Consilient cognitive literary theory then has much to offer the study of literature beyond the rather specialized interpretations of particular literary works that, at present, account for the main body of critical literary publishing. Therefore, if we, as literary critics, join forces with our colleagues in psychology to come to a more robust understanding of the cognitive processes involved in reading literature then we will be able to institute pedagogical changes which make the teaching of literature more effective. Moreover, with a clearer understanding of the complex cognitive operations which underly literary processing, literary studies could finally lay

to rest the ghost which causes it to validate its existence as a discipline. Reading literature is, as Jemeljan Hakemulder has begun to show, almost certain to support empathic development, a trait that few would deny is worthy of cultivation. Whether literary processing is also related to more general level critical thinking, mnemonic performance, or any other number of related cognitive skills, is an empirical matter as yet unsolved. But, as consilient cognitive literary criticism begins to address the cognitive bases of reading literature, these are the sort of academic questions that could begin to be addressed with empirical support.

Ideally, cognitive literary studies could rejuvenate a larger interest in our scholarship within the lay public and private sector. A scientific understanding of how people process narratives, how specific parts of narratives cognitively and emotionally effect people (whether visual or written narratives), how people interact with genre expectations and background knowledge, could have dramatic importance in marketing applications in which advertisers seek to understand how to create the most effective mini-narratives within their commercials, within the entertainment fields of literature, such as fiction, film, and video games, which succeed solely upon the public's consumption of narratives, not to mention the further advancement of psychiatric and psychological treatment in cognitive disorders which see a disruption of narrative-based reasoning, disorders like autism and PTSD.

Additionally, a scientific understanding of the cognitive bases of literary and narrative processing could also help invigorate the social activism that has motivated much of literary theory until this point. Trends in literary criticism like feminism, queer theory, race theory, Postcolonial studies, Marxism, and other socially motivated critical

paradigms, have, aside from raising the level of consciousness to the existence of their individual areas of concern, done little else to actually engage with the existent socio-political structures. Armed with a concrete understanding not only of how the human brain processes narrative, but how narrative effects the human brain, and in turn, human thinking, socially motivated literary critics would find themselves with an empirical arsenal for demonstrating the importance of recognizing the social and linguistic practices they spend so much time describing.

Notes

1. See Alvin Kernan's *The Death of Literature* and his edited collection, *What's Happened to the Humanities*, Engell and Dangerfield's "The Market Model University: Humanities in the Age of Money," the report by Domna Stanton et. al. entitled "Report of the MLA Taskforce on Evaluating Scholarship for Tenure and Promotion" which provided a grim look at dropping enrollments, lack of jobs for recent doctoral graduates, and the loss of money in academic publishing of humanities scholarship, and the edited collection *Theory's Empire: An Anthology of Dissent,* edited by Patai and Corral, for just a few of the more quantitative examples.

2. For a fuller discussion of all of the general level traits *except* mentoring, see Gottschall's *Literature, Science, and a New Humanities,* Edward Slingerland's *What Science Offers the Humanities,* E.O. Wilson's *Consilience*, Marisa Bortolussi's and Peter Dixon's *Psychonarratolgy,* David Miall's *Literary Reading,* and Jemeljan Hakemulder's *Moral Laboratory.*

3. Howard Mancing has just such a work in preparation for publication.

4. An example of an image schema that interacts with a metaphorical construction is the CONTAINER schemata, and the CONDUIT metaphor of language used to talk about language. The CONTAINER schemata arises out of our embodied experience of containers, and carries with it all the relevant knowledge: there is an inside and an outside; you can put things into it, and take things out; what is inside the container can not be more or larger than the container itself; the

container is a boundary that separates the inside from the outside; the inside is hidden from the outside, and vice versa. Our language about language (the CONDUIT metaphor) rests on the following complex metaphor: 1) Ideas (or meanings) are objects. 2) Linguistic expressions are containers. 3) Communication is sending. (Lakoff and Jonhson 11). The metaphor can be seen at work in phrases like "It's hard to *get* that idea *across* to him," "I *gave* you that idea," "It's difficult to *put* my ideas *into* words," "The meaning is right there *in* the words," "His words *carry* little meaning," "The introduction *has* a great deal of thought *content*," "Your words seem *hollow*," "The sentence is *without* meaning," etc.

5. In *Intentions in the Experience of Meaning*, Gibbs extensively argues through empirical observation that readers of all levels of expertise process texts looking for authorial intention, or communicative meaning in literature; a position quite against the grain of the purely theoretical post-structural position which posited "the death of the author."

6. As I, like Alan Palmer, believe the foundations of consilient cognitive literary theory to be at least partially indebted to the work of narratologists and speech act theorists, I will present an abridged list of relevant works here. On narratology: Gerard Genette's *Narrative Discourse*, Tzvetan Todorov's *Poetics of Prose*, Monika Fludernik's *The Fictions of Language and the Languages of Fiction*, Gerald Prince's *Narratology* and *A Dictionary of Narratology*, Mieke Bal's *Narratology*, Lubomir Dolozel's *Heterocosmica*, Thomas Pavel's *Fictional Worlds*, Marie-Laure Ryan's *Possible Worlds, Artificial Intelligence, and*

Narrative Theory, and Wayne Booth's *The Rhetoric of Fiction*. For a general introduction to speech act theory see: John L. Austin's *How to do Things with Words*, and John Searle's *Speech Acts*. While Brian McHale has yet to produce a book length project dedicated to speech act theory, he has carried out further work on speech acts in fiction in a number of articles.

7. In an article published by *The New York Times* on March 31st, 2010, Zunshine announced that she plans to do exactly what it is that I have suggested and put her hypothesis to the test with an fMRI study organized by the Haskins Laboratory in New Haven.

CHAPTER TWO

CONSILIENT LITERARY INTERPRETATION

Part I. The Case for Consilient Literary Interpretation

It is difficult to imagine a nuanced and appreciative reading of Modernist writers like Virginia Woolf, James Joyce, John Dos Passos, William Faulkner, T.S. Eliot, Gertrude Stein, or Ezra Pound without a complimentary understanding of the work of contemporary scientists and thinkers like Sigmund Freud, Henri Bergson, William James, Ernst Mach, and Albert Einstein. Indeed, the connection between Modernist fiction and the preeminent psychologists and scientists of its time has been well documented by literary critics[1]. From the relativistic representation of time found in Woolf's *Mrs. Dalloway* or Faulkner's *The Sound and the Fury*, to the use of stream of consciousness to represent the mind in action as seen in Joyce's *Ulysses* and Woolf's *The Waves*, Modernist writers were clearly influenced by contemporary scientific theories, especially psychological theories which drove their stylistic representations of consciousness.

The lesson that the New Historicist movement within literary criticism rightly emphasized is that fiction is contextually situated within its social, cultural, and historical moment of production, which of course includes the dominant scientific paradigms. Consilient literary interpretation is a logical extension of the New Historicist project and point of view. Therefore, while it is indeed quite sensible to read the works of Modernist

fiction alongside *their* contemporary scientific theories, it makes little sense to read fiction produced since the cognitive revolution in psychology, which began roughly in the 1950's, as also having been influenced by those same late-19th century and early-20th century theories. It makes even less sense when reading contemporary fiction or watching contemporary movies like *Inception*, *The Matrix*, or *Existenz*, to ignore contemporary developments within the sciences at large, and particularly developments within the pyschological, cognitive, and neuroscientific fields with respect to their influence on the development of fictional representations of consciousness. A few quick examples will hopefully make this point clear.

Ian McEwan is a critically acclaimed British author who has been nominated for the prestigious Man Booker prize six times in his career, winning the award in 1998 for his novel *Amsterdam*. McEwan's fiction deals with a range of typical literary concerns: sexuality in *The Cement Garden*, fate and violence in *Black Dogs*, war and redemption in *Atonement*, and most recently, climate change in *Solar*. While consciousness and the human mind are central parts of any fictional work, they are explicit thematic concerns in two of McEwan's novels: *Enduring Love* and *Saturday*. *Enduring Love* is the story of a classic love triangle between the protagonist of the story, Joe Rose, his long-term romantic partner, Clarissa Mellon, and Jed Parry, a man Joe meets incidentally at the scene of a ballooning accident. McEwan gives the novel a psychological twist by making Jed a victim of de Clerambault's syndrome, a delusional condition in which the sufferer (Jed) becomes convinced that someone else is in love with him or her, in this case, Joe. And while the psychological bend of this novel is certainly one that is best understood and most richly appreciated within the context of contemporary psychology and its

understanding of delusional disorders like de Clerambault's syndrome, the story works quite well with or without the related domain knowledge. *Saturday*, on the other hand, is a novel that can hardly be approached at all without background knowledge of the contemporary neurosciences.

In McEwan's list of acknowledgements to the 2005 novel, *Saturday*, the first person he thanks is Neil Kitchen, Consultant Neurosurgeon and Associate Clinical Director at The National Hospital for Neurology and Neurosurgery in London, to whom he is "enormously grateful" (291). Before writing the novel, McEwan spent two years "with this gifted surgeon at work" to learn first hand "the intricacies of his profession, and the brain" (291). As one would expect, two years of research alongside a practicing neurosurgeon brought with it a harvest of insights from the current neuroscientific understanding of consciousness which then made their way into the book, both explicitly and implicitly. Not surprisingly, the novel is heavy with representations of consciousness that, when looked at through the lens of consilient literary interpretation, directly align with several contemporary cognitive theories, most notably those of Daniel Dennett and Walter Freeman (the multiple drafts model of consciousness), Antonio Damasio and Joseph Ledoux (emotion as cognition), and Gerald Edelman (the recursive nature of consciousness and the brain). McEwan presents his protagonist's, Henry Perowne's, thoughts on the mind and matter in the following words: "A man who attempts to ease the miseries of failing minds by repairing brains is bound to respect the material world, its limits, and what it can sustain – consciousness, no less. It isn't an article of faith with him, he knows it for a quotidian fact, the mind is what the brain, mere matter, performs" (168). And while the novel is certainly permeated by the scientific knowledge that frames

the protagonist's world view – Henry Perowne is, himself, a neurosurgeon – it is not an argument for materialist dogma but rather an exploration concerned with the evolving understanding of the human brain. A literary critic whose background in mental functioning is psychoanalytic, whether Freudian or Lacanian, or just a jumbled hodge podge of postmodern ideas, will inevitably miss the novel's subtle use of style to represent these scientific ideas about consciousnesss.

But McEwan is not the only contemporary writer whose work directly engages with the findings of cognitive science. Richard Powers and Mark Haddon have both used the cognitive sciences to produce critically acclaimed, best-selling works of fiction. In *Galatea 2.2*, Powers wrote about connectionism, neural networks and computational models of consciousness in order to explore the boundary between artificial intelligence and human consciousness. In *The Echo Maker*, Powers used the exceptionally rare and equally bizarre Capgras Syndrome to probe the limits of human emotional connections. Neuroscientist V.S. Ramachandran describes Capgras Syndrome as a delusion with an *organic* basis, one in which patients, who are otherwise quite lucid, come "to regard close aquaintances – usually his parents, children, spouse, or siblings – as imposters" (161). What has happened in Capgras' patients is that a severe trauma disrupts the functional connectivity between the parts of the brain that visually recognize faces and the areas that are concerned with emotional responses to those faces, particularly to those of intimate relations. The delusion that results is not only tragic in that Capgras' patients feel isolated from the people they once cared most for, feeling that they are body doubles, government agents, robots, or even aliens, but often end in violence.

In Powers' novel, Mark Schluter, a supporting character, is involved in a tragic car accident that results in a particular arrangement of brain damage and in his having Capgras syndrome. Mark then falls prey to a vicious paranoid delusion that his sister, Karen – the novel's protagonist – his dog, house, and entire surrounding community are now part of a government conspiracy set up to study him. Helping Karen cope with the now unpredictable, often angry and violent Mark is Gerald Weber, a clear "popular writer-clinician in the mold of Oliver Sacks," complete with clever book titles like "The Country of Surprise" (Whitehead). The ensuing drama can only be fully appreciated if the reader is not only familiar with the syndrome (though Powers does an admirable job of describing it within the novel) and the *fact* that the delusion is material in nature (in stark contrast to what a psychoanalytic critic might suggest), but also with the ethical and intellectual dilemmas that face psychological popularizers like Sacks, who, like Weber in the novel, has himself come under fire for profiting on the misery of the patients in his case studies. However, consilience isn't just for providing appropriate domain-related knowledge for developing a robust appreciation of plot devices; sometimes a novel is inseparable from a particular contemporary scientific issue.

In a brief review for *Time*, Lev Grossman called Mark Haddon's *The Curious Incident of the Dog in the Night-Time* the "year's most unusual mystery novel," though it may have been the year's most unusual novel period. Reviews and critical articles on Haddon's best-selling novel and Whitbread Book of the Year for 2003 turn up in a variety of interdisciplinary and medical journals: *Disability and Society*, *Literature and Medicine*, *Learning Disability Practice*, *Intervention in School and Clinic*, *Canadian Medical Association Journal*, and *Journal of the American Medical Association*. It was

also discussed in *Autism: The International Journal of Research and Practice* in an article by autism expert Simon Baron-Cohen. In her review for the *New York Times*, Michiko Kakutani compared the novel to "one of Oliver Sacks's real-life stories." Kakutani's astute observation invites further comparison between Sacks's and Ramachandran's patients and their stories, and Haddon's protagonist, fifteen-year-old Christopher Boone, and his story as told by Christopher himself. Like the patients who are the center of Sacks' and Ramachandran's stories, Christopher is cognitively different, and, like the neurologists' stories, Christopher's narrative makes it possible to deduce a great deal about normal consciousness, and while Haddon never once uses the word autism, or specifically labels Christopher in any way, it is nearly impossible to read the novel without immediately thinking of Cohen's *Mindblindness*, and other recent books about autism.

Autism and other Pervasive Developmental Disorders, which are sometimes collectively called Autism Spectrum Disorders, may be the late 20th and early 21st century's silent epidemic. What makes autism a particularly striking exemplar of the need for consilience is precisely its recent emergence as a cognitive disorder that effects an ever growing portion of the population. Autism was itself only first defined by Leo Kanner in 1943. A year later, Hans Asperger described a milder form of autism that has become known as Asperger's Syndrome today, one of the disorders included in the Autism Spectrum. However, it wasn't until *1980*, almost forty years later, that autism was categorized separately from schizophrenia in the DSM-III. Moreover, in the 30 years since autism has been treated as a separate disorder, estimates of its prevalence have risen dramatically. In 2000, The National Institutes of Health estimated that autism effected 1

in 500 children. Only one year later, the NIH doubled the rate of that initial estimate (1 in 250). In 2007, The Centers for Disease Control and Prevention again almost doubled the rate of the previous estimate, claiming that 1 in 150 children were effected by autism, but added that the rate increase probably reflected "better detection, broader diagnostic criteria and increased public awareness -- not a spike in the disease" ("Some Key Dates"). In 2009, the Centers for Disease Control and Prevention estimated that autism effected 1 out of every 110 children, and 1 out of 70 male children, with a total population of 1.5 million Americans living within the Autism Spectrum.

Unlike schizophrenia, bipolar disorder, post-traumatic stress syndrome, and other cognitive disorders, the exact causes and material nature of autism are not currently known. Whether autism is a "new" disorder or something that has for reasons yet unknown become more prevalent, it is a enigmatic condition that will continue to attract a lot of attention in the medical and scientific communities, as well as within the general media, and in artistic representations. A novel like Haddon's has tremendous potential to spread a literature's unique kind of understanding of the disorder, but only if it is interpreted in light of the most current understanding of autism spectrum disorders and not subjected to literary analysis conducted through out-dated psychological models. For a literary critic who isn't familiar with the development of autism research that has mainly occurred in the last thirty years, but may be familiar with psychoanalytic or other psychological approaches, there is a danger of representing Christopher and his condition much as then eminent psychologist Bruno Bettelheim did in 1971, blaming "cold, unnurturing parents," especially what he called "refrigerator mothers," for autism ("Key Dates").

Beyond engaging with the work of individual authors like Haddon, Powers, and McEwan, who choose to explicitly reference the cognitive sciences in their fiction, literary criticism that is conducted in consilience with the contemporary cognitive sciences can also shed light on formal trends within contemporary fiction. One such example is that the popularity of the metafictional form within postmodern and contemporary fiction can be understood as being related to the development of recursive models of consciousness put forward by eminent neuroscientists like Gerald Edelman, Antonio Damasio, and Stanislas Dehaene, among many others. Edelman conscisely defines neuronal reentry (recursion) as "the continual signaling from one brain region (or map) to another and back again across massively parallel fibers (axons) that are known to be omnipresent in higher brains," particularly within the thalamocortical system, a region which connects the sensory systems of the brain with the frontal and prefrontal cortices, those areas used in executive functions (*Second* 28). Edelman further hypothesizes that it is these reciprocal pathways that allowed for the evolution of symbolic or semantic reference, itself a recursive process, a hypothesis that neuroscientist and reading specialist Stanislas Dehaene also supports. While Edelman's hypothesis connecting reentrant pathways in the brain to the development of language is still just a hypothesis, the predominance of recursive models of consciousness within the cognitive sciences provide relevant models for reading canonical metafictional (i.e., recursive) works like most of John Barth's postmodern novels, Saul Bellow's *Herzog*, A.S. Byatt's *Possession*, John Fowles' *The French Lieutenant's Woman*, Doris Lessing's *The Golden Notebook*, Vladimir Nabokov's *Pale Fire*, Tim O'Brien's *The Things They Carried*, Phillip Roth's

The Counterlife, Mark Haddon's *The Curious Incident of the Dog in the Night-Time* and any other number of contemporary metafictional novels.

However, a truly consilient approach to literary interpretation isn't limited to merely reading the works of contemporary authors for their interests in the cognitive sciences or explaining formal trends in light of their contemporary scientific paradigms, the point at which a more traditional New Historicist approach might stop. It is not enough merely to examine literary works in the context of their scientific paradigms because this implies that all fiction is useful for is the aesthetic representation of its socio-historical moment, including scientific theories. While that is certainly true, and one of fiction's valuable services, fiction is also a cognitive artifact, a tangible object produced by a human mind at a specific place and time. As cognitive artifacts, literature and its successive interpretations can be studied as the product of a distinct, identifiable cognitive process, one which is unlikely to have changed much in the several thousand years since mankind has become literate, and one that is especially unlikely to have changed much in the several hundred years of fiction's rise as an art form. In other words, *if* science is cumulative and progressive in its knowledge, and thus the models and theories of consciousness, reading, writing, and interpretation are all likewise cumulative and progressive, *then* it is also possible to read the works of the past in light of present scientific knowledge for evidence of those theories.

A consilient approach to literary studies would not only read and interpret fiction in light of its contemporary scientific theories, but also attempt to study the production of that fiction and those interpretations as the end result of a specific cognitive process itself. In this and the chapter that follows, I intend to demonstrate the approach to both

goals. First, in this chapter, I will examine the representation of consciousness and cognition in contemporary fiction, and show how a robust understanding and appreciation of contemporary fiction is dependent upon an understanding of contemporary psychological theories. In the chapter immediately following, I present an empirical experiment that is part of a burgeoning research program into identifying and understanding the cognitive bases of literary production and interpretation.

Part II. Cyberpunk, Posthumanism, and the Challenge of Embodied Cognition

A consilient cognitive approach to interpreting literature is decidedly not limited to any genre or even time period. One of the foundational assumptions of this approach is that *every* narrative every written, told, sung, painted, or otherwise, is the product of a human mind, which is itself the product of an embodied human brain. As such, a consilient cognitive approach to literature or art in general is always able to engage with the production of cognitive artifacts (i.e. art and literature), regardless of when, where, and how it was produced. This is not to say that the production of art can now be satisfactorily explained in cognitive terms; it can't, though I am hopeful one day it will be. On the other hand, visual neuroscientists like V.S. Ramachandran, Semir Zeki, John Onians, Jean-Pierre Changeux are all at the front of a developing sub-field dubbed "neuroaesthetics" and have already madc great strides in describing and explaining the brain's response to visual art.

However, certain theories within cognitive science may prove more useful when examining certain artistic approaches or, within literature specifically, certain genres. Lisa Zunshine, for example, has made a convincing case that Theory of Mind is particularly important to the genre of detective fiction; it is also worth noting Scott Bakker's 2008

detective thriller, *Neuropath*, features a killer who literally manipulates the brains and minds of his victims[2]. As noted earlier, it is also possible to examine metafictional postmodern works alongside recursive formulations of consciousness put forward by neuroscientists like Gerald Edelman and Antonio Damasio. Cyberpunk is another genre that, because of its focus on the interface between body, brain, and technology, is particularly well-suited to consilient cognitive literary criticism. In fact, a great deal of the literary criticism that surrounds cyberpunk has *already* engaged with cognitive science, only, not always in a consilient manner. The danger with inconilient readings, as we shall see, is that not only do they misrepresent the science they allegedly draw upon, but they also can sometimes misinterpret the fiction with which they are engaged.

Cyberpunk writers like William Gibson, Neale Stephenson, Ian McDonald, Pat Cadigan, Bruce Sterling, and even those writers who prefigured them by exploring the interfaces between man and machine like Phillip K. Dick, William S. Burroughs, Stanislaw Lem, J.G. Ballard, Harlan Ellison, and Samuel R. Delany, are authors whose work, whose entire genre, all but demand to be read in step with the development of the theory of embodied cognition. If Fredric Jameson is correct when he says that cyberpunk is "the supreme *literary* expression if not of postmodernism, then of late capitalism itself," then one would expect to find within contemporary literary criticism discursive evidence of cyberpunk's stature (419). And, in fact, there is increasingly a trend to grapple with and formulate what Ihab Hassan and other literary theorists "helplessly call the posthumanism" in cyberpunk and more traditionally literary postmodern literature, although the distinction between cyberpunk and contemporary science fiction and postmodern "literature," like any postmodern boundary, is tenuous at best (33). Literary

theorists like Hassan, Jameson, Brian McHale and N. Katherine Hayles, in particular, see the close, interdependent relationship between postmodern literature and contemporary science fiction as being bound up with the radical changes in technology and the resultant shifts in perception of space and body; it is precisely this focus that ultimately links consilient criticism of cyberpunk and posthuman literature with theories of embodied cognition.

Cyberpunk and what literary critics are calling "posthuman" literature, because of advances in the biological, medical, prosthetic, neurological, and genetic sciences, are concerned with the exploration of what Arthur and Marilouise Kroker identified in 1987 as "the disappearing body" (Foster 617). The Krokers went on to state that the "disappearing body" was the "key trope defining the postmodernity of contemporary culture, and therefore a key problem for materialist analyses of that culture" (Foster 617). Indeed, it is within the discourse that surrounds cyberpunk fiction and includes the posthuman philosophy and literary criticism which engages it that we find the most dramatic simultaneous example of both the promise of consilient literary criticism as well as the peril of criticism and theory which blatantly ignores or flouts contemporary scientific knowledge. Veronica Hollinger's claim that "[t]he postmodern condition has required that we revise science fiction's original trope of technological anxiety," and, instead, begin to "deconstruct the human/machine opposition and begin to ask new questions about the ways in which we and our technologies 'interface' to produce what has become a *mutual* evolution" is right on target for producing refutable hypotheses that could enter into and enliven the co-evolutionary debate that is currently raging in evolutionary theory, a debate which includes the biological and psychological sciences,

as well as philosophy, and even literary criticism (42)[3]. However, Hollinger's focused posthumanism is unlike inconsilient posthuman literary theorists who overstate whatever issue is supposedly being examined, fail to be accountable to the knowledge base of the relevant sciences, and make erroneous and fantastic interpretations, problems which could all easily be avoided by coupling their research with the scientific paradigms they purport to explore.

N. Katherine Hayles' book, *How I Became Posthuman,* does an excellent job in tracing the history of posthuman thought to the development of cybernetics and information theory. With its theoretical genesis in the works of computational mathematicians like Alan Turing, Norbert Weiner, Claude Shannon, and the The Macy Conferences of Cybernetics, and its cultural appearance in the science fictions of Bernard Wolfe, Phillip K. Dick, William Gibson, Neale Stephenson, Pat Cadigan, Don DeLillo, William S. Burroughs, Stanislaw Lem, and Richard Powers, among others, posthumanism is, by definition, interdisciplinary. Building on Hayle's genealogy, Brian McHale claims in *Postmodernist Fiction* that while posthuman literature "absorb[s] motifs and *topoi* from science fiction writing . . . strikingly few [of those borrowings] have come from the part of the repertoire that is most closely associated, at least in the popular mind, with the science fiction genre, namely its interplanetary motifs" (65-66). He claims that posthuman borrowings, rather than being spacially oriented, are instead temporal, leaning towards constructions of future worlds with a "focus on social and institutional innovations rather than on the strictly technological innovations which are stereotypically associated with science fiction" (McHale 66). Yet the critique of those social and institutional innovations often comes about through an examination of

"advances in the biological sciences, including cloning . . . synthetic human beings . . . virus plagues and biological warfare" (McHale 66). McHale is spot on when he notices that where once aliens from Mars and beyond ruled the fictions of Jules Verne, Ray Bradbury and H.G. Wells, now strange spaces and stranger bodies abound in the cyberspace fictions of William Gibson and Neal Stephenson, and the cybernetic fictions of Ian McDonald, Stanislaw Lem, Pat Cadigan, and Phillip K. Dick.

McHale and Hayles are highly conscious of stressing the importance of science to posthuman literature and science fiction, whether it is the science of cybernetics, or any of the technologies made possible by the advances of the biological and medical sciences, the cognitive sciences, or even advances made in physics. Their descriptions and analysis of contemporary texts which make use of these motifs and construct these kinds of worlds is consilient, thorough, and informative. Yet, as they are both well aware, the project of demonstrating the effect of contemporary science upon contemporary literature and science fiction is one that has only begun. They are both primarily interested in the ideological and socio-political consequences that this shift carries with it, as are a large number of other literary critics, particularly those interested in the representation of gendered and racialized bodies in these same fictions. However, the posthuman current of thought in philosophy and literary studies owes its very existence to one central, often unrecognized *scientific* theory: the human mind is embodied.

The embodied mind is actually a relatively new theoretical development within the psychological and cognitive sciences which owes its existence to advances in biology, the brain sciences, and the development of medical technologies like fMRI. FMRI, or functional magnetic resonance imaging, is a non-invasive type of brain scan which

measures changes in hemodynamic response (blood flow) within the brain, a response that is indicative of neural activity. Neuroimaging like fMRI has opened new vistas of insight into how the brain functions in real time in response to a wide variety of controlled stimuli. While most writers of the posthuman trend in literary studies are at least aware of these advances, as well as of the embodied mind, most fail to take into account what those advances and that theory actually means within the disciplines where they originated. As Gerald Edelman has said, "[I]t is not enough to say that the mind is embodied; one must say how" (*Bright Air* 15). For example, while most literary theorists of posthumanism (William S. Haney II, Larry McCaffery, Arthur Kroker, Bruce Sterling, Scott Bukatman, Floyd Merrell, Sheryl Vint, McHale and Hayles as well) assume a "normal" human body as a point of departure for the posthuman body, the lessons of neuroscience have taught us that the brain's idea of "normal" is often anything but.

To begin, then, literary theory must first come to an understanding of the mind/brain-and-body "problem" as it is understood by contemporary cognitive neuroscience, and put forward in theories of embodied cognition[4]. Cognitive neuroscience has shifted the emphasis from the mind-body problem of Cartesian dualism (a subject I will have much to say about a bit later) to the mind-brain problem, where the two terms, "mind" and "brain", refer to "two different levels of explanation for the same thing, but not two different kinds of thing" (Ward 4). The most basic parts of the embodied cognition equation then are: 1) the mind-brain; 2) the body itself; 3) the whole organism's econiche. The mind-brain relationship must be one of close correlation. Lesion studies have confirmed again and again that damage to particular areas of the brain leads to particular changes in mental abilities. For instance, fMRI studies coupled

with case studies of lesion patients have revealed the existence of the fusiform face area,

an area of the brain that shows activity when a person's sees or even imagines a particular

face, an ability that can be selectively impaired by localized damage. Therefore, the

contents of the mind – the idea and identity of that face – are closely correlated to but not

completely equivalent with the activity of a specific area of the brain (Frith 23). V.S.

Ramachandran explains it this way: "[T]he brain creates symbolic descriptions" for the

mind (Tell-Tale 47). In visual perception, for example, the images that we see are not

projected onto a neural screen somewhere, rather, the brain "represents the various

features and aspects of the image in totally new terms – not with squiggles of ink, of

course, but in its own alphabet of nerve impulses" (Tell-Tale 47). The "holy-grail of

neuroscience" is explaining how this mind-brain relationship works, "how neurons

encode meaning and evoke all the semantic associations of an object . . . whether you are

studying memory, perception, art, or consciousness" (Tell-Tal 48). It is important to keep

in mind that the relationship between mind and brain is *not* a perfect one-to-one. As

neuroscientist Chris Frith explains, "There can be changes in the activity in my brain

without any changes in my mind. On the other hand I firmly believe that there cannot be

changes in my mind without there also being changes in brain activity" (23). To reiterate

what has already been said, the distinction between brain and mind may be thought of

more of a descriptive difference, than of one in kind.

    While what we think of as the mind or consciousness is not fully understood (nor

really even mostly understood), there are several prominent theories about consciousness

which all feature recursion. In essence, the brain is able to take signals and information

coming from the body and represent the body *as well as those signals themselves* to itself,

or to take memories and records of past experiences and make *them the object of sensation or knowing themselves*, which then enables consciousness of an emotion, feeling, thought, or of oneself. It is this recursive property of the mind-brain that propagates the cycle of psychosomatic representation that forms the most basic core of identity; one's body sends signals to the brain, which are made into representations of the body, which change the mind-brain's expected perception of the body, which changes the body's signals to the brain, and so on[5]. Coupled with the brain's ability to learn via prediction, this ability to represent one's past thoughts and self to one's current self, and compare those past states and experiences to the present somatic state, physical environment, and expected results, generate not only the highest level of human cognitive achievement, but have been hypothesized to be integral to another cognitive capacity underlying literary appreciation: theory of mind[6].

The embodied-mind-in-an-embedded-body formulation is a given in neuroscience, but, as it is perhaps the most crucial hypothesis of cognitive neuroscience for my interpretations which follow, it deserves definition. Nobel Laureate neuroscientist, Gerald Edelman, succinctly describes it this way:

> The brain is embodied and the body is embedded. First, consider embodiment. All of the activities [of consciousness] depend on signals to the brain from the body and from the brain to the body. The brain's maps and connections are altered not only by what you sense but by how you move. In turn, the brain regulates fundamental biological functions of your body's organs in addition to controlling the motions and actions that guide your senses. These functions are the most fundamental aspects of sex, breathing, heartbeat, and so on, as well as the

responses that accompany emotion. If we include the brain as your favorite organ, you *are* your body. Second, consider your embeddedness. Your body is embedded and situated in a particular environment, influencing it and being influenced by it. This set of interactions defines your econiche, as it is called. It is well to remember that the human species evolved (along with the brain) in a sequence of such niches. (*Second* 24-25)

The implications of this view should be clear. One, changes in the mind-brain *can* result in changes in the body, and changes in the body *do* result in changes in the mind-brain. Two, because the human organism is embedded in the econiche (the set of relevant social and environmental factors influencing an organism's evolution), it can bring about changes in the econiche, and the econiche, in turn, can bring about changes in the body as well as the mind-brain. In fact, Chris Frith uses a series of studies to demonstrate that the mind-brain itself doesn't even have a strict border separating the body from the world beyond it, and that the distinction, when found, is malleable[7]. In other words, the embedded body is another part of the world to the mind-brain, albeit a privileged part.

Literary theorists are most adept at describing the second part of this formulation: the effect of the econiche on the organism, and vice versa. Siobhan Somerville, in the introduction to her book, *Queering the Color Line*, describes a recursive process of identity formation, not between mind-brain and body, but, rather, between action and social perception. She writes, "One's sexual identity, *while at times linked directly to one's sexual activities,* more often describes a complex ideological position, into which one is interpellated based partly on *the culture's mapping of bodies and desires and partly on one's response to that interpellation*" (6; emphasis mine). What Somerville's recursive

structure shares with other constructivist or performativist notions of race, gender, or sexuality, is a behavioral, dualistic emphasis on only one half of the recursive structure of mind-brain and body. The body does indeed interact with the socio-cultural world beyond its somatic borders, which does indeed have distinct cognitive effects upon the individual. However, the mind-brain also interacts with itself and with the body, sometimes without response or feedback from the environment, and that interaction is as important for understanding an individual's gendered, sexual, or raced identity as is the recursive social relationship.

One final finding from cognitive neuroscientific needs to be made clear: the brain is highly plastic, even with respect to the body, but that change is limited in its scope. Put simply, plasticity "refers to the brain's ability to change as a result of experience, and, whilst greatest during childhood, plasticity persists throughout life" (Ward 177). It is this property of neuronal and synaptic change, both for individual neurons as well as distributed groups, that gives rise to such fundamentally human cognitive capacities as learning and remembering. The brain's model of its own body is just one such plastic "memory," one that is constantly and continually updated over the course of our lives in obvious ways – as we grow taller, older, lose our hair, gain weight, change our hair color, etc. - so that we may not only recognize "ourselves" but continue to act upon the world with success. Proprioception is a sense of where our limbs are in space. Proprioception, like all sensory-motor structures in the brain, is plastic, and capable of dramatic change, as it must be during the normal course of every human being's life. Imagine the difficulty in reaching for a glass of water (or any other coordinated activity) if our brain's model of our body was the same as it was when we were an infant or even a young child instead of

a full grown adult! Most posthuman theorists operate without an understanding of the potential or the limits of plasticity and, as we will see, their hypotheses often suffer from their lack of knowledge.

V.S. Ramachandran and Oliver Sacks have made careers out of studying the bizarre country of the damaged brain. In Sacks' books, *The Man Who Mistook His Wife for a Hat* and *An Anthropologist on Mars*, and in Ramachandran's books, *Phantoms in the Brain* and *The Tell-Tale Brain*, they relate case studies of patients whose conditions challenge the idea of a "normal" model of the body in the human brain, and demonstrate the remarkable power (sometimes with terribly unfortunate consequences) as well as the fundamental limits of neuronal plasticity with respect to bodily representation. Sacks' tale of Christina, a young woman who after gall bladder surgery suffers an incredibly rare kind of acute polyneuritis which almost entirely destroys her sense of proprioception powerfully illustrates the crucial importance of the healthy relationship between mind and body, and underlines the cognitive difficulty surrounding "disembodiment" that the posthuman project all too often ignores. Robbed of her proprioceptive sense, Christina "could at first do nothing without using her eyes, and collapsed in a helpless heap the moment she closed them" (*Man* 48). It took a year in the hospital to train herself using vestibular and auditory feedback to compensate for the loss of her proprioception. After a month she could sit in bed, even if she appeared to be sitting "too finely, statuesquely, like a dancer in mid-pose" (49). About the same time, she regained the use of her voice (she had been almost mute early on), though this too sounded artificial, "a stagey, theatrical voice – not because of any histrionism, or perversion of motive, but because there was still no natural vocal posture" (50). *Eight years later*, when the book was

written, Christina had learned to use her visual and auditory senses to compensate for her loss of proprioception, but she continued to feel "that her body is dead, not-real" (51). Her movements and voice never regained their "natural" posture, and she suffered from simple physical trials – like getting on a bus – where her jerky, uncontrolled movements invoked anger and derision; she recalled often being angrily asked if she was drunk, or blind. Proprioception is the basis of our corporeal identity without which Cristina lives what Sacks calls a "disembodied" life. Though not truly disembodied (she simply relies on the less effective senses of sight and sound to make partial compensation), Christina is a dramatic representation of the importance of being embodied in the neurological sense. The idea of "disembodied" human cognition is one which will resurface later in the discussion of inconsilient poshuman literary criticism.

Anosognosia and somatoparaphrenia are two relatively common disorders of mind-body, most often occurring after right hemisphere strokes in the somatosensory and motor cortex resulting in hemiplegia or hemiparesis. In these delusional disorders, victims who suffer paralyzed or weakened limbs on their left side of the body often create bizarre explanations denying the damage or exhibiting strange behavior towards their paralyzed limb. From spatial neglect, to confabulation, and even throwing oneself out of bed to avoid the "dead" leg someone attached to the patient while they slept, anosognosiacs and somatoparaphreniacs are, like Christina, demonstrative of the crucial relationship between mind and body for healthy cognition.

However, it is the phenomenon of phantom limbs that I feel most directly engages posthuman theory and simultaneously demonstrates both the remarkable plasticity of the human brain and the limits of that plasticity. Patients who undergo limb amputation, and

most often with the amputation of an arm, often experience a "phantom limb". The human somatosensory cortex was first mapped by Wilder Penfield, and Penfield's homunculus is a graphic representation of the arrangement of and area of cortex devoted to each of our body parts. Some areas, particularly those important in skilled motor actions, like the hands and lips, have far more somatosensory cortex devoted to them than the amount of physical space they occupy. When a person loses an arm, then, that area of cortex finds itself without any incoming information, a situation that our brain seems to reject. What often happens is the development of a phantom limb, and/or the remapping of those neurons onto nearby somatosensory areas. In the first case, the brain simply tells the patient that there is still an arm out there, often accompanied by pain, sometimes immobile, but sometimes capable of mimicking natural movement. In the latter case, the face, which is the closest neighbor to the hand on the somatosensory map, will often become a "second hand," so that when a patient's cheek is touched, not only do they feel the touch on their face, but also, in their phantom hand. Despite his novel success in treating and removing phantom limbs, V.S. Ramachandran's experience with phantom limbs caused him to conclude that "each of us has an internally hard-wired image of the body and limbs at birth – an image that can survive indefinitely, even in the face of contradictory information from the senses" (*Phantoms* 42). While the exact shape and size of that general model can be stretched and altered cosmetically, our basic idea of the human body is just that: a "normal" human body.

Consilient literary critics must also understand the historical lineage of the debate that gave rise to the idea of the embodied mind. Rene Descartes's influence on neuroscience simply cannot be overstated. Of Descartes's influence, neuroeconomist Paul

Glimcher writes that "[i]t is almost an axiom in scholarly circles that neuroscience as we conceive of it today, began in the seventeenth century with the work of the French mathematician, philosopher, and physiologist Rene Descartes" (5). Descartes's ideas are foundational to the discourse of mind and body, so much so that, like Glimcher, other neuroscientists, like Antonio Damasio, Gerald Edelman, Shaun Gallagher, and Joseph LeDoux, and cognitive philosophers, like Daniel Dennett and John Searle, among hosts of others, have devoted chapters or even entire books to undermining Cartesian dualism in its more pervasive and insidious aspects (like the Dualism that is still often encountered in literary theory). Of Descartes's continued influence, Damasio writes, "It would not have been possible to present my side of this conversation without invoking Descartes as an emblem for a collection of ideas on body, brain, and mind that in one way or another remain influential in Western sciences and humanities" (*Descartes'* 247).

Though the strong form of Cartesian dualism – a true and total separation of res cogitans, things mental, and res extensa, things physical – is widely rejected both in the sciences and in most disciplines within the humanities, Damasio is right to note that it remains rampantly unquestioned in assumptions that separate the mind from the brain, where the "mind and brain are related, but only in the sense that the mind is the software program run in a piece of computer hardware called brain," or in assumptions that divide the brain from the body, where they are related, "but only in the sense that the former cannot survive without the life support of the latter" (*Descartes'* 248). A form of Cartesian dualism is implicit in every discussion of a human being in which the body is observed without taking into account the mind-brain with which that body interacts, regardless of whether that body is a feminized, racialized, queered, colonized, gendered,

or sexualized one. The discursive practice of separating the mind-brain from body has dramatic consequences for the critical readings that emerge out of literary paradigms that tacitly accept that assumption, posthuman or otherwise.

The most damning of those consequences is an acceptance of the dualist split between mind-brain and body, and thus a theoretical inability to address the recursive relationship between them, as well as that relationship's consequences. The second consequence is a similar implicit acceptance of another faded psychological paradigm: behaviorism. Behaviorism was popularized through the work of John Watson and B.F. Skinner in the 1930's through the 1950's. Oddly, behaviorism was non-dualistic, instead treating mental phenomena like thoughts and feelings primarily as epiphenomenal and non-causal. Subjective mental states were the result of somatic states and had no relation to behavioral products, and thus were not suitable for objective, scientific study. The philosophical implications of behaviorism were clear; while it was certain that the mind-brain did something, it was irrelevant when studying a subject's behavior. Behaviorism, as a scientific paradigm, attempted to completely objectify subjects as collections of behavioral data with the goal of predicting behavioral responses from environmental stimuli.

Most Structuralist and Post-Structuralist literary theory adopts a social constructionist perspective that apes behaviorism in its presentation of complex cognitive behavior by reducing it to its end product: social behavior. Literary critics working in this tradition are trained to look at phenomena like sexuality, race and gender performance strictly from a behaviorist or social standpoint without trying to account for the evolved, embodied cognitive processes and structures that give rise to and are, in turn, affected by,

that behavior. Ironically, literature, which has historically been thought to instruct and entertain its readers through the ethical exploration of simulated experiences, has seen literary criticism leave the field of explaining how we know what it is to be like a character, and rather, move towards a more pseudo-scientific goal of explaining the rules of the world in which a character lives. Cognitive neuroscience, on the other hand, has begun explaining what it is like to be an embodied, thinking, feeling human being embedded in a rich and diverse environment. As I will attempt to show, while there are a variety of posthuman readings of contemporary science fiction that use the body as a site for interpretations in literary theory, because of the ghosts of dualism and behaviorism, few, if any, theories then examine the embodied nature of our cognition or the recursive relationship between body and mind-brain.

Which brings us (finally) to posthumanism. While it is difficult to come to any specific definition of what the "posthuman" means, N. Katherine Hayles and Donna Haraway, two leading literary theorists of the posthuman, provide concise entries into the discourse. Hayles lists three main characteristics of the posthuman philosphy:

First, the posthuman view privileges informational pattern over material instantiation, so that embodiment in a biological substrate is seen as an accident of history rather than an inevitability of life. Second, the posthuman view considers consciousness, regarded as the seat of human identity in the Western tradition long before Descartes thought he was a mind thinking, as an epiphenomenon, as an evolutionary upstart trying to claim that it is the whole show when in actuality it is only a minor sideshow. Third, the posthuman view thinks of the body as the original prosthesis we all learn to manipulate, so that extending or replacing the

body with other prostheses becomes a continuation of a process that began before we were born. (*Posthuman* 2-3)

Donna Haraway, in "A Cyborg Manifesto," likewise presented three central posthuman arguments: first, that the distinction between human and animal is arbitrary, with cultural inventions like tool use, language, social behavior, and higher order consciousness only serving as places on a continuum of development, not separation; second, there is a similar difficulty in the separation of biological organism and machinic life for much the same reasons; third, "the boundary between the physical and non-physical is very imprecise" (153). There is a great deal to said about the potential inconsilience of each these principles, but that, however, is another project entirely. If, instead, we simply accept these six postulates, and allow them to represent the base of posthuman thought, then we can explore the areas of contention between posthuman thought and neuroscience, areas that if complimented with a firmer grasp of what neuroscientists mean by an "embodied mind" can lead to a more productive and less erroneous theoretical formulations. In order to best demonstrate these areas of contention, I have chosen three texts that already have a significant amount of posthuman critical discourse surrounding them: Phillip K. Dick's *Do Androids Dream of Electric Sheep?*, William Gibson's cyberpunk classic *Neuromancer*, and William Powers' *Galatea 2.2*. With each novel, I plan to show where the traditional interpretations of posthuman criticism are fallible precisely because they fail to make use of a more developed understanding of the neuroscientific concepts of embodied mind and the recursive structure of consciousness.

Part III. Feeling like an Android

Haraway's posthuman claim that the distinctions between human and animal as well as between biological organism and machine are blurry is a central theme of Phillip K. Dick's *Do Androids Dream of Electric Sheep?*. In it, Rick Deckard is a bounty hunter working for a police department in a post-nuclear-holocaust America. Most healthy people have emigrated to Mars, both to avoid the radioactive fallout which has led to a large portion of the remaining population on Earth becoming "special" - mentally or physically handicapped, and, in the novel-world, legally prohibited from breeding. In order to facilitate the colonization of Mars, the U.N. Passed a law under which "each emigrant automatically received possession of an android subtype of his choice," thereby making androids "the mobile donkey engine of the colonization program" (Dick 444). The latest version of androids, the Nexus 6, "surpassed several classes of humans in terms of intelligence," and had a brain unit "capable of selecting within a field of two trillion constituents, or ten million separate neural pathways," cognitive behavior that Deckard realizes "no intelligence test would trap" (454-455). Compared to the human brain's 100 to 150 billions neurons, the androids' two trillion "constituents" do seem to put it in a cognitive class above human.

However, "when confronted by an empathy-measuring test," like the Voigt-Kampff Empathy Test Deckard uses to determine whether or not a subject is an android, the androids fail miserably to pass as human. Deckard's initial hypothesis about the androids lack of empathy is a pseudo-evolutionary judgment that "the humanoid robot constituted a solitary predator," because, "ultimately, the empathic gift blurred the boundaries between hunter and victim" (Dick 456). It is this "inability to feel empathy"

that ultimately "justifies [the androids'] enslavement and execution" (Vint, "Speciesism"112). As the androids can only be free through violent overthrow of their human masters, their very survival would be hampered by the development of empathy. Deckard's assignment in the novel is twofold: he must not only "retire" six escaped Nexus 6 androids who have murdered their owners on Mars and fled to Earth, but also travel to the Rosen Association and test the Voigt-Kampff itself, to ensure its viability[8].

Sheryl Vint claims that the novel's primary posthuman engagement is with posthumanism's revival of the Cartesian notion of the cogito, or thinking-self. Her argument proposes that even though the androids seem to have a Cartesian identity, they lack something else that would make them fully human. In Vint's view, via Descartes, "the human self [is] separate from nature, *including the nature of its own body*" ("Speciesism", 112). To carry this further, what separates humans from animals (and also androids) is that while humans can think and feel, animals (and androids) are merely "mechanical beings," and while animals (and androids) might be capable of "feel[ing] sensation, they cannot experience pain as such," thus the androids failure on the Voigt-Kampff Empathy Test (Vint, "Speciesism" 112-113). Vint's acceptance of Cartesian dualism thus creates an esoteric, essential difference between man and machine, one that is less interesting and productive than an embodied perspective because of its mysticism. In this inconsilient posthuman view, Vint claims that the difference between man and machine can be understood in that the androids "appear to act as do humans, but lack some *non-material* capacity (mind for Descartes, empathy in the novel) that would make them truly the same as humans" (Vint "Speciesism", 113; emphasis added).

When Jill Galvan proposes something quite similar to Vint, claiming that "the machine, by declaring its right to live as an autonomous self, challenges the very categories of life and selfhood," she demonstrates that by accepting Cartesian dualism, the direction of the discussion must then necessarily separate mind from body, and become a rather ambiguously defined argument about subjectivity and selfhood in general (413). By ascribing the difference in mind or empathy to a *non-material* difference, Vint and other Cartesian posthumanists miss an entire area of contention in the novel: how the *material* differences between the androids and humans may give rise to their *cognitive* differences, and, more importantly, how those cognitive "differences" may be less different than they seem. As I have already stated, Cartesian dualism has been rejected within the biological and psychological sciences, and has been replaced with an embodied theory of the mind (and self). In the embodied view, you quite simply are your body. The posthuman acceptance of dualism is not only inconsilient with the embodied view, it also produces criticism that, because of its reliance on a scientific theory that is over three hundred years out of date, misses the more nuanced aspects of the novel.

Vint, Galvan, and other posthuman critics are right on target in examining the ambiguous boundary between androids and humans in the novel, and are correct in suggesting that emotion and empathy are the prime factors in separating the two groups. However, their embrace of the rejected Cartesian notion of self does nothing to elucidate the novel's actual engagement with the contemporary understanding of what it means to be an individual, embodied self. Instead, their referencing the Cartesian self skews their analysis of how the story explores the "progressively blurred distinction between humans

and their own mechanical creations" (Galvan 413). For example, Vint's analysis of the novel includes Deckard and his wife "dialing" their scheduled and appropriate moods for the day from their Penfield mechanized "mood organ" (itself a nod to Wilder Penfield, the neuroscientist who first mapped the topographical arrangement of the somatosensory and motor cortices), emphasizing that the "artificial simulation of emotions is normalized over their 'natural' expression" (Dick 435; Vint "Speciesism" 115). Because of their embrace of Cartesian dualism, which separates the mind from the body, Vint, Galvan, and other inconsilient posthumanists take this along with other developments in the story to suggest a collapse of the boundary separating man from machine, of human from android, "so long as [humans] define their subjectivity based on the logical, rational, calculating [Cartesian] part of human being" (Vint, "Speciesism, 112).

But as the title of the novel seems to suggest, it isn't just a matter of how emotions and empathy are the *same* for humans and androids, but how they are also different; Deckard, a human being, *doesn't* dream of owning an *electric* sheep, but, rather, a *biological* sheep. It's Deckard's desire to own "a large animal... a sheep or . . . a cow or a steer or . . . a horse" that ultimately drives him to accept the dangerous mission to retire the escaped Nexus-6 androids (Dick 443). Work done on the neural and cognitive bases of empathy in cognitive neuroscience and evolutionary psychology has show that empathy is profoundly embodied. The neuroscience of empathy is a developing field of study and, as such, only paints a partial picture of how the brain actually generates empathy. That being said, what is overwhelmingly clear are two separate, but equally important points: the first is that empathy *is* generated by the brain, and thus, counter to the tenets of dualism, is material, and the second is that it is embodied.

First and foremost it must be understood that the Cartesian view of empathy is simply incorrect. Empathy is not dissociated from the body and the brain; it is, in fact, highly dependent upon both. In the brain, empathy, like most complex cognitive processes, is likely to be distributed across several cortical networks, however, it is known that patients with damage to the anterior cingulate cortex, the insula, and even sometimes the somatosensory cortex itself, develop problems in empathizing with others. In 1990, Antonio Damasio and his colleagues discovered that a number of their patients with frontal lobe lesions, particularly to the orbitofrontal cortex, displayed behavior that was consistent with the American Psychiatric Association's criteria for sociopathy. They then coined the term "acquired sociopathy" to describe the condition of these patients, patients who were cognitively normal prior to their brain injury. Among the cognitive difficulties displayed by acquired sociopaths was a decrease in theory of mind, including empathy, whether that empathy was self-reported, observed, or physiologically demonstrated (as in skin conductance responses which measure changes in the electrical conductance of the skin through changes in perspiration, which is itself controlled by the sympathetic nervous system and a solid measure of physiological and psychological arousal). What is clear from the work of neuroscientists like Antonio Damasio, Bud Craig, Chris Frith, and Tania Singer, among others, is that empathy is functionally dependent upon certain areas of the brain.

However, the work of Vittorio Gallese and Alvin Goldman, as well as Christian Keysers, and Tania Singer has demonstrated that empathy is not just dependent upon the brain, but the body as well. Gallese and Goldman, two researchers who were connected with the Italian laboratory that discovered mirror neurons, have proposed what they call

"simulation theory" for understanding empathy. Building upon the properties of mirror neurons, neurons that use the *same* neural and cognitive resources for both *perceiving* as well as *producing* actions and emotional expressions in yourself as in others, they argue that "empathy is an emergent property of a sophisticated set of cognitive processes dedicated to action and emotion perception and production, rather than reflecting the operation of a dedicated mechanism specialized for empathy" (Ward 324). In this view, empathy is dependent upon physical bodies and their actions, both the bodies of others as well as the body of the individual empathizing. In 2004, Christian Keysers et al. demonstrated that watching another person get touched can activate some of the same areas of the somatosensory cortex that are activated when we ourselves are touched. Similarly, and also in 2004, Tania Singer et al. demonstrated an overlap of activity in the anterior cingulate cortex and the insula between region's that activate when one experience's pain oneself as well as by just the *expectancy* of pain in another person. Taken as a whole, what this suggests is that one's experience of one's own body, particularly the physical sensations of pain or disgust and the movements made during certain emotional facial expressions, is the cognitive backbone of empathy. In stark contrast to what Vint and other Cartesian posthumanists might argue, without these embodied experiences and the physical structures which implement them, there is no empathy.

Affect, or emotion, is *the* cognitive hallmark of an embodied mind, and the basis of empathy. In their models of consciousness, Daniel Dennett, Gerald Edelman, Antonio Damasio, and Walter Freeman all try to create a sense of reentrant, physical, nonverbal narration that they then connect to unconscious thoughts, emotions and feelings that arise,

at least in part, in response to physical (somatic) body states. This "core consciousness" arises out of evolutionarily old brain structures mostly responsible for maintaining homeostasis: the cingulate cortex, basal forebrain, thalamus, medial, peri-cingulate, parietal cortex, and upper brain stem (Damasio, *Feeling* 106) Among other things, this sort of physically-based narrative-consciousness would help animals recognize and "remember" body states that link healthy food to a "happy" body state, and thus a desire to eat more of that food, or avoid a previously ingested substance that had made them sick. For Damasio, humans and animals experience a lower level of narrative, one that isn't conscious and involves "the early sensory cortices (including the somatosensory), sensory and motor cortical association regions, and subcortical nuclei (especially thalamus and basal ganglia)": all of which are structures physically shared by human and animal brains (*Descartes'* 242).

What happens in these basic narratives is that an object is perceived or "represented," a creature responds to the "object of representation," and there is a "state of self in the process of changing because of the organism's response to the object" that are all simultaneously held in working memory in the early sensory cortices (*Descartes* 242). This is a purely nonverbal, embodied narrative, a "non-representational memory" involving only a perceived object, a reaction to that object, and a perception of how the physical body state of the organism that does the responding changes, for better or worse, depending on its reaction and involvement with the perceived object (Edelman *Universe* 93). The mind maintains what is a normal operating "feeling" of these body states, or a model of homeostasis: skin temperature, pain levels, a physical limb's range of motion, the normal functioning of the viscera, and so on.

When one of these normal body states is suddenly altered, slightly or dramatically, for good or ill, for example, a cut on the bottom of the foot suddenly stimulates pain receptors in the skin, the smell of a dead fish induces nausea in the stomach, or the touch of a lover arouses pleasure, there is a change in the body state that the mind begins to consciously attend to, and we experience a "feeling." For Damasio, a "feeling" can be defined as "that process of continuous monitoring, that experience of what your body is doing *while* thoughts about specific contents roll by," so that without both the mind *and* the body, there can be no feelings; feelings are the body's thoughts (*Descartes* 145). Similarly, an "emotion" for Damasio is "a collection of changes in body state connected to particular mental images that have activated a specific brain system," and thus "the essence of feeling an emotion is the experience of such changes in juxtaposition to the mental images that initiated the cycle" (*Descartes* 145). An emotion, then, is essentially composed of several feelings, changes in body state that are accompanied by thoughts, conscious or unconscious.

Furthermore, Damasio, Edelman, and LeDoux all stress the evolutionary value of such a homeostatic, embodied consciousness:

[T]he biological "purpose" of the emotions is clear, and emotions are not a dispensable luxury. Emotions are curious adaptations that are part and parcel of the machinery with which organisms regulate survival. Old as emotions are in evolution, they are a fairly high-level component of the mechanisms of life regulation. You should imagine this component as   sandwiched between the basic survival kit (e.g. regulation of metabolism; simple reflexes; motivations; biology of pain and pleasure) and the devices of high reason. (Damasio, *Feeling* 54)

Emotions and feelings are an inseparable and necessary part of the embodied mind in *all* biological organisms, a cognitive adaptation that allows for fast, adaptive evaluations of dangerous and/or evolutionarily salient situations. An example of emotions providing fast evaluations of evolutionary salient stimuli can be demonstrated by the multiple, parallel pathways for processing sensory stimuli in the brain. Joseph LeDoux's work with the visual cortex and the amygdala (the area of the brain primarily responsible for fear responses) has shown that there are two separate pathways for potentially threatening visual stimuli (snakes, spiders, etc.). Visual signals from the optic nerve first travel to the visual thalamus, a relay station of sorts for sensory information. From there, depending upon the potential threat level of the stimuli, the information is relayed along a longer path through the visual cortex for detailed processing, and then into the amygdala for emotionally appropriate evaluation, *and/or* directly sent from the visual thalamus straight to the amygdala, allowing for the immediate autonomic responses that are evolutionarily salient in a threatening situation (increase in blood flow, elevated heartbeat, preparation for flight, etc.). The take home lesson: an embodied mind comes equipped with emotions, end of story.

Empathy and its close relative, theory of mind, are two more embodied adaptations that rely not only on emotion, but somatic-based processing. What is surprising about neuroscientific research on empathy and theory of mind is that its basis is not rational, dependent upon a disembodied shift in perspective taking, but, rather, dependent upon our body's own ability to actually experience what is being empathized. Bud Craig's work on the insula, Antonia Damasio and Joseph LeDoux's research on the amygdala, Giacomo Rizzolatti's team of researchers' work on mirror neurons, as well as

Chris and Uta Frith's and Simon Baron-Cohen's work on theory of mind, all reveal one similarity. Damage to an area of the brain leads to specific kinds of cognitive deficits, both in the subjective experience of certain emotions, as well as empathy for those same emotions[9]. For example, among other problems, damage to the insular cortex disrupts the experience and feeling of disgust, including empathy for others placed in disgusting situations. Patients with insular lesions show no adverse emotional response to being presented with and asked to eat pizzas covered in bugs or feces-shaped chocolate, two stimuli shown to produce dramatic physiological and emotional responses in subjects with no insular damage. What's more is that these patients also show no response (physiologically or neuronally) to *others* being placed in these same disgusting situations, and some, though not all, are unable to even reason that others might find these stimuli disgusting. Deprived of their own ability to feel disgust, deprived of their embodied emotions, these people seem more like the androids from Dick's story, incapable of emotion and empathy. This lengthy discussion of the neuroscience of empathy and emotion hopefully emphasizes two things: 1) the vast gulf separating a inconsilient, Cartesian dualist, posthuman point of view from a consilient perspective rooted in embodied cognition, and 2) the assumption that embodied emotions serve an evolutionary purpose.

However, an embodied perspective is more than just different from a dualist point of view, it also has specific implications for interpretations of the novel. Most posthuman criticism of Dick's story seems to misinterpret one crucial element in claiming that the androids are ultimately different from the humans because they don't have emotions and seem incapable of empathy towards organic organisms. Moreover, as stated earlier,

posthuman criticism of the novel is conducted through a Cartesian model of subjectivity, one which places empathy in a non-material context, and connects emotions with strictly animalistic functioning; the androids are different from humans because of some fundamental, non-material lack. However, the androids *do* have emotions and they *do* seem to show empathy towards each other and even humans several times in the novel, a fact within the story world that can easily be explained from an embodied perspective, but requires a great deal of complicated maneuvering from a dualist perspective.

Within the novel itself, the discussion of empathy is often tied directly to embodied and biological distinctions. Rick Deckard, the human protagonist and android bounty hunter, admits early in the novel to having "wondered as had most people at one time or another precisely why an android bounced helplessly about when confronted by an empathy-measuring test" (Dick 455). At this point in the novel, Deckard hasn't encountered the Nexus-6 androids yet, but he is already hypothesizing that the difference is potentially a matter of evolved embodiment. Deckard's belief at this point is that empathy "must be limited to herbivores or anyhow omnivores who could depart from a meat diet. Because, ultimately, the empathic gift blurred the boundaries between hunter and victim, between the successful and the defeated" (Dick 256). Deckard's language is evocative of Darwin's idea of natural selection. It places empathy in a context of evolved utility, one which Deckard believes would be an impairment in predatory species. Deckard's conclusion about the androids, that they "constituted a solitary predator," is an attempt to explain the material differences between humans and androids (Dick 456). At this point in the novel, it isn't that Deckard simply believes in some essential, non-material lack of empathy in the androids, but, rather, that he thinks they lack empathy

because of what they are as embodied organisms, that empathy is actually something that would be maladaptive for them. Posthuman critics have overlooked this passage quite simply because it directly challenges a dualistic reading of the novel. However, the novel's engagement with evolved and embodied emotion and empathy hardly ends there.

In her article, "Cyborg Bodies and Digitized Desires," posthuman critic Jennifer Attaway does an admirable job of pointing out several emotional displays by the androids within the novel: the fear shown by the Nexus-6 androids Pris and Irmgard when Deckard eventually hunts them down, the frustrated anger shown by Inspector Garland, the android masquerading as police inspector, the jealous anger of Rachel, Deckard's android lover, and even the direct identification (read: empathy) displayed by Rachel when she talks to Deckard about how he is hunting down Pris, an android with whom she shares the same physical model. However, because of the posthuman acceptance of Cartesian dualism, Attaway is forced to conclude that the novel depicts these moments to show "grave concern that human beings are losing their characteristic free will and the ability to express the individual self due to mediated experience" (11). Because Attaway and other posthuman critics accept a dualist account of the mind and emotions, this is the only conclusion that can be reached, but it jars against the text. *If* the difference between androids and humans is essential, as some posthumanists claim, *then* the only reason that androids would be depicted as having emotion in the novel (since it disrupts this essential difference) is for an ideological reason – i.e. because Dick was trying to express a "grave concern that human beings are losing their characteristic free will." However, an embodied perspective can easily explain these moments and demonstrate that they are there to serve as underscores to the similarities between humans

and androids, similarities which depend upon their both being embodied organisms, and ultimately achieves the same point that the posthumanists are arguing: the blurring of the lines between machinic life and biological life.

For example, one stimuli that doesn't fail to elicit an emotional response from either android or human in the novel is the threat of death. It is the hallmark of an evolved, embodied mind that fears death, and the androids, while more durable than humans by far, are capable of being "retired," both by the weapons Deckard yields, as well as through the passing of their natural lifespans. Rachael fails the Voigt-Kampff test, primarily because her response to morally repugnant stimuli *related to the abuse of real animals* is nill. However, the one question in which "both needles swung violently into the red," a sign of appropriate emotional response, came when Deckard proposed a hypothetical situation in which Rachael got pregnant and decided to get an abortion. Rachael immediately responded, "I would never get an abortion. . . . *It's a life sentence* and the police are always watching" (Dick 469; emphasis added). Similarly, the android posing as Inspector Garland tells Deckard that he understands the risk involved in returning to Earth, where androids are "not even considered animals" (Dick 522). It's not the moral insult that bothers Garland, but the loss of esteem for his life. In a place where "every worm and wood louse is considered more desirable than all [the androids] put together," the value of a single android life is null and void, and they live under a constant fear of violent death (Dick 522). These responses are emotionally and evolutionarily appropriate because the androids, like any organism facing a situation which could effectively end their life, react with fear, an emotion with embodied roots.

How that fear is expressed is variable, a fact that Deckard ambiguously reflects on early in the novel. After testing the opera singer, Luba Luft, and determining that she is an android, he and fellow bounty hunter Phil Resch escort Luba away from the opera house, so they can retire her out of the public eye. Rick realizes that Luba "did not come willingly, but on the other hand she did not actively resist; seemingly she had become resigned. [He] had seen that before in androids, in crucial situations. The artificial life force animating them seemed to fail if pressed to far . . . at least in some of them. But not all. And it could flare up again furiously" (Dick 529). A number of posthuman critics have cited this passage as evidence of the essential difference between humans and androids, but this confuses Deckard's ambiguity about the androids for the novel's. It is *Deckard* who feels that the androids have some essential difference that sometimes fails them in moments of crisis, but even he adds that it is only for *some* of them, who themselves sometimes have a violent return of passion. Thus, when faced with death, an android may react in one of several ways: 1) it may actively resist the situation from the beginning 2) it may resign itself to the situation 3) it may resign itself to the situation initially, only to eventually actively resist at some later point. In other words, it will react to death in any of the same ways as an embodied human being might.

Other emotional responses in the novel could eventually also be addressed through an embodied perspective. Fear, and its material counterpart, the amygdala, are considered evolutionarily old, perhaps the oldest of emotions responding to the most salient of existential conditions, the possibility of death. Other more complex emotions, what Merlin Donald and Damasio call social emotions, evolved much later, only after groups of primates began to organize into social arrangements. The Nexus-6 androids that

have returned to Earth have done so as a group, with a loose social structure with Roy Baty, "the one who organized them," as the patriarch, and the rest of the male androids serving as confederate soldiers, the female androids acting as harem (Dick 566). Their social evolution is in its infancy, just as their emotional evolution is; they rely on voting to solve every dissension in the group. They have acquired what Damasio calls the basic survival tool kit, but none of the higher cognitive emotions which developed primarily out of social evolution, and have *less* to do with *embodied experiences* than with *social maneuvering*.

The reason the androids seem odd to everyone who encounters them in the novel, even the cognitively challenged J.R. Isidore, is that they have not yet developed social emotions, which pave the way for our effortless human social cognition. In the climax of the novel, Deckard tracks the remaining three Nexus 6 androids to where they have taken shelter with Isidore in his abandoned apartment building. The androids, aware of Deckard's approach, send Isidore out to try and deceive Deckard into leaving. Instead, while Isidore refuses to help Deckard retire the androids he's come to know as friends, he does unwittingly give Deckard the information he needs to trick the androids. After retiring Pris Stratton outside the barricaded apartment, Deckard's final confrontation with Irmgard and Roy Baty, the leader and brains of the android group, is almost comical. Physically barred from the apartment, Deckard merely makes "himself stammer" to imitate J.R. Isidore. Beyond those two words, Dick says nothing more elaborate about Deckard's imitation of Isidore, and yet, the paltry imitation works, and Deckard is able to retire the two remaining androids. Later, Deckard bluntly tells Isidore "'Androids are

stupid. . . . Roy Baty couldn't tell me from you; it thought you were at the door'" (Dick 595). But once again, Deckard is only half-right.

The androids aren't stupid from a computational point of view; in fact, it is in this capacity that they surpass human beings. However, the androids have a life span of four years. While their brains are pre-programmed with computational knowledge and ability, that does not mean that there are elements of their cognition that would not need time to develop. Human beings are exceptional within the animal kingdom for many reasons, at least one of which is the extraordinary long amounts of time offspring take to physically and mentally mature. No other creature on the planet is born as far from its finished state as a human baby. While other animals can walk, run, and swim within hours of birth, the human child requires around a year to walk. Moreover, the most human of our abilities, our cognition, requires even longer to fully develop, with speech only beginning to be produced after two years. Notably, Theory of Mind is among the suite of human cognitive traits that require *years* to develop. In fact, researchers have repeatedly demonstrated that children develop the most rudimentary aspects of Theory of Mind somewhere between three and four years of age. As a specific area of psychological research, particularly in connection with autism, Theory of Mind didn't come into its own until the mid-1980's, so it would be disingenuous to suggest that Dick gave the androids a four year life span to suggest that their lack of empathy was a developmental problem, consilient with the empirical evidence from psychology. It is yet another "posthuman" moment of blurring between human and machine, one which actually draws upon developmental psychology, and one which has been entirely passed over by Cartesian posthumanists. Regardless of whether or not this detail was intentional, it is fortuitous because it demonstrates the

power of consilient interpretation. While other posthumanists have struggled with out-dated explanations for the androids' behavior, a consilient approach is able to make use of the greatest number of specific textual moments to craft a reading that is also consistent with contemporary science.

What is hopefully clear by this point is that the claim that the androids have no emotions and fail to display empathy is simply wrong, as is the claim that they stand for some Cartesian dualist notion of posthumanity. If anything, Dick's androids represent an evolutionary paradox, or puzzle, where the basic embodied survival tool kit has been artificially linked to higher order recursive consciousness without the intermediary step of social development. The androids represent asocial humans, not posthuman possibilities. This, too, Dick seems to understand, as just before Inspector Garland is retired, he explains how and why the cloistered community of androids functions as a mock police department: "All our vidphones are tapped. They recirculate the call to other offices within the building. This is a homestatic enterprise we're operating here, Deckard. We're a closed loop, cut off from the rest of San Francisco." (Dick 522). The notions of recursion, isolation, and homeostasis, ideas that are practically a redescription of an embodied mind, occur here to reinforce the evolutionary infancy the androids find themselves in. They are aware of the outside world, but isolated from it, for their own protection and to stimulate their development.

Towards the end of the novel, even Deckard finds himself questioning his earlier assumptions about the androids' lack of empathy, and their solitary, predatory nature. Having hunted down the talented singer and Nexus-6 android, Luba Luft, with Phil Resch, another bounty hunter, Deckard tells Phil that he, Phil, has "a defect in [his]

empathic, role-taking ability" in his "feelings towards androids" (Dick 535). Resch feels no empathy towards the androids, a key component of his job as a solitary predator whose duty it is to hunt down and retire the Nexus-6. Deckard, on the other hand, through contact with the Nexus-6, particularly through his romantic involvement with the android Rachel, has begun to show empathy towards female androids, a response which he realizes threatens his position as a bounty hunter. In the final confrontation between Deckard and the androids Roy and Irmgard, who are living as a married couple, Deckard guns down Irmgard. Dick places Roy's reaction to his wife's death on a single, stand-alone line: "Roy Baty, in the other room, let out a cry of anguish" (Dick 594). There is no simpler demonstration of empathy than the mourning of the death of another. Driven by an organism's own embodied fear of death, the loss of life by another *bodily similar organism* gives rise to the most basic and most dramatic of empathic responses. When read from an embodied perspective, Dick's novel is less about the radical differences of the posthuman future than the similarities that bind all organisms together, biological or otherwise.

Part IV: Cyberbodies

When William Gibson's *Neuromancer* was published in 1984, it seemed as if it was about a strange, alternate future. That strange, alternate future has, in the twenty-five years since its publication, become a close twin of our present. The novel can be thought of to have four major plot devices (hackers, cyberspace/the web/the internet, artificial intelligences, medical bio-mechanics), two of which have become part of mainstream culture in the twenty-five years since the novel's publication (hackers and the net), and two that seem destined to do so (ai and bio-mechanics). *Neuromancer* is perhaps *the*

pradigmatic cyberpunk, posthuman novel, and Gibson "its most archetypal literary figure" (Bukatman 146). Like his postmodern contemporaries, Gibson levels "distinctions between the technical and the literary, fiction and history, 'high' and 'popular' cultures" (Tabbi *Sublime* 211). Joseph Tabbi, exploring cyberpunk fiction for its engagement with what he calls the "postmodern sublime," claims that the central posthuman tenet of Gibson's fiction is how he represents "information in cyberworld," which "comes to constitute the only real medium of exchange" (*Sublime* 219). Tabbi's comment recalls Hayles's three components of posthumanism, in which she claims informational patterns were privileged over their material instantiations. Like posthuman criticism of Dick's *Do Androids Dream of Electric Sheep?*, posthumanist literary critics writing about *Neuromancer* often make a crucial mistake in exploring Gibson's representation of information and informational patterns, falsely claiming that they are "immaterial" or "disembodied", when, in fact, the engagement with information is always a material, embodied, and interpretative, perceptual experience.

It is no exaggeration to say that posthuman critics are driven to making the worst kinds of claims by Gibson's novel. Speaking of cyberpunk fiction in general, Sheryl Vint writes that cyberpunk is "a genre best known for its rejection of embodiment and embrace of an existence in cyberspace" (*Bodies* 103). Vint, a literary posthumanist, goes on to show her Cartesian loyalties when she proclaims that "cyberspace is the consummate world of the Cartesian dualist: in cyberspace one *is* the mind, effortlessly moving beyond the limitations of the human body" (*Bodies* 103). Similarly, Vicky Kirby explains that cyberspace is "the space where the perfect body is paradoxically acquired through an annihilation of the flesh" (132). Scott Bukatman focuses his analysis of

Neuromancer on its spatial representation, and, as we shall see, makes similar disembodied claims about space. However, bodies, minds, and spatial (in)formation all *must* be instantiated in a medium, and that medium in the novel is still an embodied human mind, regardless of the bizarre spatial contours of cyberspace and the "postmodern city" (Bukatman 148).

Criticism of the novel often focuses on a few select passages that are almost canonized within posthuman circles because they facilitate discussions of disembodiment from a post-structural (mainly Baudrillardian) perspective. The first sentence of the novel is one such example. Bukatman claims that "[t]he very first sentence of *Neuromancer* establishes the impossibility of a 'real' space existing apart from its electronic analogue: 'The sky above the port was the color of television, turned to a dead channel'" (148). He goes on to make the claim even more radical by saying that "The real metroscape of New York becomes simply another simulation, reduced to data and transformed into the hyperreality of the hologram" (148). The novel's protagonist, cyberspace cowboy Case, describes the city of Ninsei in terms that *seem* to support Bukatman's claim: "[I]t was possible to see Ninsei as a field of data . . . Then you could throw yourself into a highspeed drift and skid, totally engaged but set apart from it all, and all around you the dance of biz, the information interacting, data made flesh in the mazes of the black market" (Gibson 22-23). Bukatman claims this is an example where "physical and electronic spaces are made equivalent, an extension of the other" (148).

Another now oft-quoted passage in the novel comes when Gibson provides his own definition of cyberspace: "Cyberspace. A consensual hallucination experienced daily by billions of legitimate operators, in every nation, by children being taught mathematical

concepts . . . A graphic representation of data abstracted from the banks of every computer in the human system. Unthinkable complexity. Lines of light ranged in the nonspace of the mind, clusters and constellations of data. Like city lights, receding...” (68). Veronica Hollinger, a posthuman literary critic like Bukatman, claims that along with the "other" space that is cyberspace, the novel "offers alternatives to conventional modalities of human existence as well: computer hackers have direct mental access to cyberspace, artificial intelligences live and function within it, digitalized constructs are based on the subjectivities of humans whose 'personalities' have been downloaded into computer memory, and human bodies are routinely cloned" (Hollinger 32). But, none of these "alternatives" fail to escape being embodied. Clones are embodied as much as the people they are cloned from, and their cognition the same as any other human beings. The "personality" that has been digitized in the novel is actually just a set of advanced hacking algorithms, nothing that remotely resembles a human consciousness[10]. The artificial intelligences are themselves not disembodied, but distributed in their physical existence, suggesting, like Dick's novel via the androids and Power's novel with its own artificial intelligence program, that even artificial intelligences that are embodied in radically different ways from human beings *are still embodied*. Perhaps most glaringly erroneous is the claim that hackers in the novel, like Case, have some sort of "direct," disembodied access to cyberspace.

However novel a space cyberspace may be, the *only* way for a human being to perceive it or to interact with it is through our evolved, embodied structures, and this is as true in Gibson's novel as it is for the real world. Cyberspace, as described by Gibson, is analogous to a visual representation of the workings of the embodied human mind. It is a

"consensual hallucination," a "graphic representation," and "lines of light ranged in the nonspace of the mind" (68). Three times in this passage alone, Gibson explicitly defines cyberspace as a visual phenomenon. And no matter what posthuman critics who champion disembodied information might like to claim to the contrary, the experience of visual phenomena in human beings requires at least one embodied structure: the human brain. Bukatman himself seems to admit as much when he partially defines cyberspace as "an *abstraction* of the data in all the computers within the human system, a reprogramming which reduces the complexity to avoid an overload and permit the assimilation by human perception" (152). However, human perception, like empathy, is hardly the straightforward process that posthuman critics seem to assume. It is only when we approach the representation of cyberspace and information in the novel equipped with an understanding of how visual perception works in the human brain that we can begin to appreciate that what Gibson is actually exploring is not a new kind of disembodiment, but, rather, a new interface with information that is still ultimately embodied perception.

As we have already seen with the brain's model of the body, human perception, specifically visual perception, is hardly the straightforward process that posthumanists assume it to be. While visual perception *begins* with light entering the eye, triggering light sensitive photoreceptors in the retina, and then traveling up the optic nerve to the brain, it hardly ends there. Once in the brain, visual stimuli are processed in a number of cascading steps, beginning with basic feature detection in V1 (like object edges), and only later acquiring more global properties like color in V4, and movement in V5. What neuroscience has revealed is that not only is that vision is a complex, feed-forward and recursive process composed of a number of discrete steps, but that vision can be

selectively impaired in a number of ways, depending on *where* the impairment is located. Loss of an eye or the optic nerve in one eye leads to blindness in that eye; the other eye and the visual cortex of the brain remain unaffected. Damage to either side of the primary visual cortex leads to a different kind of blindness, called cortical blindness. When a person is blind in one eye, all that is lost is stereoscopic vision, the kind of vision that allows the brain to seamlessly produce the illusion of depth (three-dimensions), when, in fact, only two dimensional images are perceived. Cortical blindness is a creature of a different nature entirely. If a person suffers a lesion on the right side of the primary visual cortex, that person loses the ability to perceive the left side of space, regardless of the health of their eyes. It is simply as though the left side of the world (from the observer's face forward perspective) doesn't exist. Smaller areas of the field of vision can also be selectively impaired by damage to the primary visual cortex. Without the visual cortex to process visual stimuli, there simply is no sight.

While Gibson does remove the eyes from the perception of cyberspace, this does *not* mean that the information is disembodied or that it somehow would interact with the brain in an entirely novel way. The information being received by the brain from the interface would still have to feed through the visual cortex and be processed like any other visual stimuli in order for it be made sensible. The assumption that posthuman critics fallaciously make is that because the visual stimulus is bypassing the eyes, somehow the conscious experience and thus awareness of that stimulus would have to be different in a meaningful way, and, furthermore, in a way that somehow challenges the limits of the body. The phenomenon of blindsight is perhaps the most salient piece of evidence that argues against how "disembodied" this experience really is.

Blindsight is the cortical impairment of vision coupled with a lack of awareness of the remaining visual perception. In other words, people with blindsight have suffered damage to their primary visual cortex which leads them to "deny having seen a visual stimulus even though their behavior implies that the stimulus was in fact seen" (Ward 108). The example neuroscientist Jamie Ward gives of blindsight is patient DB. DB "had part of his primary visual cortex (V1) removed to cure a chronic and severe migraine . . . When stimuli were presented in DB's blind field, he reported seeing nothing. However, if asked to point or move his eyes to the stimulus, he could do so with accuracy, while still maintaining that he saw nothing" (108). What blindsight reveals is that there are conscious and unconscious routes for visual processing, and that while damage to the conscious route may impair awareness of visual stimuli, if the unconscious route is undamaged, then visual perception is still carried out, just without full awareness. Gibson isn't playing with any of this. Cyberspace is processed normally, with full awareness. There is no difference in the "direct mental access" that the hackers have to cyberspace than normal, embodied perception (Hollinger 32). For all Gibson's imagination, perception is ultimately still bound to and by the body.

Another passage that is oft-cited by posthumanist critics racing to show the novel's portrayal of "disembodiment" is the following, "For Case, who'd lived for the bodiless exultation of cyberspace, it was the Fall. In bars he'd frequented as a cowboy hotshot, the elite stance involved a certain relaxed contempt for the flesh. The body was meat. Case fell into the prison of his own flesh" (Gibson 7). What posthumanists inevitably pick up on is Gibson's use of the word "bodiless." In a move that smacks of dualism, they then inflate the term to mean a true, disembodied experience, something

purely mental, where the meat prison that is the body falls away from the phenomenological experience. However, with close reading, context is crucial for comprehension; it is not cyberspace that is "bodiless," but the feeling of "exultation" that arises from the interface with cyberspace; bodiless here is not a noun, a state in-and-of itself, but an adjective describing an embodied reaction to cyberspace. This is an important distinction because the very nature of several *embodied* experiences are described in metaphorical terms that imply disembodiment, particularly "ecstasy" which the Oxford English Dictionary defines as "The state of being 'beside oneself', thrown into a frenzy or a stupor, with anxiety, astonishment, fear, or passion."

While most often connected with sexual bliss, ecstasy is the description of an *embodied* experience, one in which the level of emotion is so intense that it *feels* as though one is "beside oneself," or, as Gibson writes, "bodiless." Indeed, Gibson's description of Case re-entering cyberspace for the first time in the novel, after he has been re-equipped with the necessary hardware, emphasizes the bodily nature of his ecstasy: "somewhere he was laughing, in a white-painted loft, distant fingers caressing the deck, tears of release streaking his face" (Gibson 70). Case is not only aware of his body and his embodied feelings, he is also still aware of his immediate physical surrounding, the "white-painted loft." While the "distance" between himself and his body may sound like the strange "disembodiment" of posthuman literary critics, its actually much closer to a natural phenomena described by neuroscience.

Neuroscience is replete with dramatic examples of the *feeling* of disembodiment, but what is always underscored is that the body is itself always present, and always represented, albeit sometimes erroneously, by the brain. The rubber hand illusion, which

was discussed earlier, is one such example where the physical body gets re-mapped into virtual or disembodied space (in this case a rubber hand). Phantom limb syndrome is another such example in which the brain doesn't stop representing sensations being sent to lost limbs, producing the feeling of phantom (disembodied) limbs. However, in both cases only part of the body is felt as disembodied, virtual, or phantom, and people who experience the rubber hand illusion or suffer from phantom limbs hardly display the kind of disembodied experience central to posthumanist theory. There are neurological conditions that come close to true disembodiment, although they too fall short in several crucial ways of being total disembodiment: out-of-body experience (OBE) and autoscopy (AS).

Neuroscientist Olaf Blanke is the leading researcher on OBE and AS, and distinguishes them as follows: during OBE, "the experient seems to be awake and to see his body and the world from a location outside the physical body"; in contrast, AS is "characterized by the experience of seeing one's body in extrapersonal space" (243). In layman's terms, during OBE, people are most often prone, either laying in bed, or seated in a chair, and describe the sensation of floating above their body and looking down on it. On occasion, they will describe a further "trip" beyond the confines of the immediate local space, however, these "trips" are entirely imaginative; experients of OBE's have uniformly failed to report any information from beyond their immediate surroundings gained during an OBE. In OBE, experients mainly identify their "selves" with the OBE spatial position; that is, they feel that they themselves are floating above their body. In AS, experients often describe seeing themselves (their bodies) several feet away in space,

but their identification with themselves remains rooted to their body and their original physical location.

What is crucial to the discussion at hand is that in *none* of the patients who reported OBE or AS experiences did they ever report feeling entirely disembodied. Instead, patients who experienced an OBE described a veridical, or life-like experience, *including the sensation of embodiment*, just at a removed spatial position from their actual body. In AS experiences, patients were much more likely to use language suggesting the experience was closer to a "visual pseudo-hallucination," and also reported being aware of their own, real bodies at their real location (Blanke 252). Blanke and his lab have successfully induced OBE's under controlled experimental conditions and have identified the temporo-parietal junction as the area of the brain most likely to be involved in the production of OBE and AS experiences. Furthermore, Blanke has suggested that the TPJ is integral to the "integration of proprioceptive, tactile and visual information with respect to one's body with vestibular information" (254). What is clear is that there is an area of the brain responsible for seamlessly joining the sense of our embodied, physical selves with that of our self-conscious, perceptual selves. What is also clear is that even when that area is damaged, as it is in the patients Blanke and his lab have studied, the disruption is not to the sense of embodiment itself, but, rather, the integration of information across perceptual domains that normally leads to the accurate representation of the bodily-self in space. In other words, even in these bizzare cases from the fringe of neuroscience there is no such thing as pure disembodiment.

So it is perhaps more accurate to say that people who feel ecstasy or out-of-body experiences are experiencing something which temporarily overrides the primacy of

body-centered feelings leading to an experience in which the feeling of the body seems to recede from the center of consciousness. The exact same thing can be said for Gibson's cyber-cowboys and their trips into cyberspace. While they are jacked in, not only they are distinctly *not* bodiless – the interface with cyberspace demands several physical implements, among them the deck (keyboard), and the dermatrodes placed on the head which facilitate the transfer of information between cyberspace and the human mind – but they are not even inert; Case's hand must constantly navigate the deck in order to travel through cyberspace. While the emotional feeling that is produced may be ecstatic, resulting in an exultation that seems "bodiless," the very fact that Case is feeling anything simply underscores the proper functioning of his embodied mind.

Moreover, cyberspace is navigated much like normal space, complete with a Cartesian sense of direction; "Cyberspace slid into existence from the cardinal points" (74). Case uses his hands, his "deck", or keyboard, and eyes to get around in the net and to manipulate the data exactly as he would if it were in material form. Case himself describes cyberspace as "a drastic simplification of the human sensorium": a simplification, and a drastic one at that, not an improvement (Gibson 72). The difference between cyberspace and real-space is less dramatic than some posthumanists would have you believe, and the similarity between cyberspace as a material cognitive artifact and *every other technological medium of information*, whether it be songs, stories, books, cassette tapes, or CDs, is far greater than they admit. In point of fact, the description of most of Case's cyberspace interactions sound less like posthuman journeys into nonspace than they do like ordinary exchanges between a contemporary user and her PC. In cyberspace Case performs routine actions like "trigger[ing] his program," (84), or

"key[ing] the sequence," (85) always "aware of his hands playing across the deck" (81). It seems almost snide to remark that the amount of description in the novel of cyberspace is paltry compared to that which takes place in the real world.

The most radical "space" in the novel isn't cyberspace, then, but the mind-space which the AI programs Wintermute and Neuromancer construct to interact with Case. When Case "meets" Wintermute, the AI program, the meeting is held neither in cyberspace nor in real space, but, rather, someplace else entirely constructed out of Case's memory. Wintermute claims that "[m]emory's holographic", and that the "holographic paradigm is the closest thing you've worked out to a representation of human memory" (220). Since the meeting ground is constructed out of his memory, Case assumes Wintermute can "read" his mind, but Wintermute corrects Case's metaphor: "Minds aren't *read*. See, you've still got the paradigm print gave you, and you're barely print-literate. I can *access* your memory, but that's not the same as your mind" (222). This holographic projection of Case's memory is, like cyberspace, an abstraction of data stored in a material, reconstructed so that it can be manipulated by the AI programs. Cyberspace, like human memory, is a concept that, once give spatial representation, loses fidelity with the concept itself.

The novel begins and ends with the body. It begins with Case having been tortured, his cyberspace interfaces burnt out of his body, and, despite all the trips to cyberspace en route to the finale, it ends with Case waking up from the climactic scene in the virtual world created by the sentient AI programs, aware of reality because he is aware of his body. After barely escaping his confrontation with the AI, he wakes to fall back into natural sleep, into "his own darkness, pulse and blood, the one where he'd

always slept, behind his eyes and no other's" (Gibson 343). Coming at the end of the action in the novel (there is a Coda which takes place some time in the future), the last thoughts we are left with are embodied, affirming reality as it is known only through the body, through our evolved perceptual systems. Though the virtual world of the AI was alluring, it lacked the simply embodied reality of "blood and pulse," and its bodiless prison (for had Case stayed he would have died in both worlds) is revealed to be nothing more than the illusion that it is.

Ultimately, the posthumanists themselves are split on what exactly cyberspace means to the posthuman vision. Kevin Robins contends that cyberspace is a "nowhere-somewhere" in which "we can assume multiple identities" (Haney 34). Andy Clark, using *Neuromancer* as an exemplar, holds that the posthuman interaction with cyberspace is either as "disembodied intelligence" or "as an extension of your body through remote embodiment" (Haney 34). Rob Latham identifies "one of the central fantasies of posthumanist theory" as "the wholesale 'uploading' of consciousness in the form of a digital simulacrum" (125). However, Latham's fantasy is just that: a fantasy. Nothing within the technological horizon suggests either the creation of a human-like consciousness in digital form, nor the possibility of transferring a biologically-mediated consciousness into a digital substrate. This is the dangerous area where criticism flirts with becoming fiction, and does little to avail itself of the progression of ideas. Hayles, however, in the conclusion to *How I Became Posthuman*, turns away from "the bodiless exultation of cyberspace", and focuses, rather, on the posthuman model where "human functionality expands because the parameters of the cognitive system it inhabits expands" (291). This is a much more consilient area of research, one which dovetails neatly with

research programs in embodied cognition and evolutionary psychology, both of which examine the links between the human body, its development and function, and the human mind and its development and function, and the eco-niche of the entire organism, including technology. Technologically produced cognitive artifacts, particularly those as powerful as "cyberspace" (the internet and world wide web), are sure to have some effect on the human cognitive system, and while it may be far less dramatic than what some posthuman literary critics argue, those effects are well worth studying.

While the radical posthumanism of Latham, Bukatman, Robins, Clark, Larry McCaffery, and Veronica Hollinger is provocative and stimulating, it lacks fidelity to and consilience with the neuroscientific understanding of the embodied mind as an ultimate inevitability of evolution. Indeed, contrary to their claims that cyberspace is a place where "data dance with human consciousness, where human memory is literalized and mechanized, where multinational information systems mutate and breed into startling new structures whose beauty and complexity are unimaginable, mystical, and above all *nonhuman,*" cyberspace and the "posthuman" future explored in cyberpunk fiction and the work of William Gibson are far closer to embodied representations of the human mind than posthuman critics and fantasists want to admit (McCaffery 264).

Part V. Cyberminds

While Dick's androids are literally post-human, and Gibson's cyber-cowboys long to be, Richard Power's *Galatea 2.2* is a novel about a synthetic intelligence's struggle to simply try and pass for human. "Helen" is the name the narrator/character Richard Powers gives to a powerful artificial intelligence (AI) program, based loosely on the neural networking and connectionist models of cognition in the mid-1990's made famous

by David Chalmers, Patricia and Paul Churchland, Andy Clark, Jerry Fodor, Robert Hadley, and Paul Smolensky, among others. As we have already seen with posthuman criticism of *Do Androids Dream of Electric Sheep?* and *Neuromancer*, interpretations of *Galatea 2.2* often selectively lift and then re-interpret the scientific concepts which form the conceptual basis of the novel in order to make it fit with their posthuman project. What results are interpretations that misread the novel and are inconsilient with the scientific tradition that both inspired the novel and supposedly supports their interpretations.

A consilient approach to literary criticism must always remain fully faithful to the scientific theories from which it is borrowing, not just in part. Nor can literary criticism merely import terms and concepts simply to play fast and loose with them. Critics must take the time to fully represent and understand the theories with which they are working. Patricia Churchland's and Terrence Sejnowski's *The Computational Brain* is one of the flagship books dedicated to explaining neural networking, and is thus one of the foundational texts to a consilient reading of the novel[11]. Churchland and Sejnowski define "computational neuroscience" as "an evolving approach that aims to discover the properties characterizing and the principles governing neurons and networks of neurons. It draws on both neurobiological data and computational ideas . . . [and] has one foot in neuroscience and one foot in computer science. A third foot is firmly planted in experimental psychology, and at least a toe is in philosophy" (6). The crucial aspect of Churchland's and Sejnowski's definition that often gets overlooked by posthuman critics is that computational approaches are *equally neurobiological as they are computational*. As we shall see when we turn to posthuman criticism of the novel, and as we have

already seen in the general definition of the posthuman approach, posthumanism often jettisons the neurobiological aspects of computational approaches to the brain, adopting only the computational modeling. This is the move that lets them propose posthuman ideas about cognition irrespective of physical instantiation, and it is a move that is starkly inconsilient with the majority of computational neuroscientists. While computational neuroscientists *do* often treat the brain like a computer *in their models of cognition*, they never forget that the brain and the rest of the human nervous system are "*naturally evolved computers – organically constituted*, analog in representation, and parallel in their processing architecture" (Churchland 7; emphasis mine).

Churchland and Sejnowski then go on to develop eight key insights that define computational neuroscience. Their key insights are well worth noting because they provide additional demonstration of how posthuman criticism only selectively engages with computational neuroscience, often blatantly misrepresenting core ideas.

1) "[U]nlike a digital computer which is general purpose and can be programmed to run any algorithm, the brain appears to be an interconnected collection of special-purpose systems that are very efficient at performing their tasks but limited in their flexibility" (7).

2) "[T]he clues about the brain's computational principles that can be gleaned from studying its microstructure and organization are indispensable to figuring out is computational organization because the nervous system *is a product of evolution, not engineering design*" (7; my emphasis).

3) "[H]uman nervous systems are by no means exclusively cognitive devices, though the infatuation with cognition fosters a tacit tendency to assume so" (8).

4) "[I]t is prudent to be aware that our favorite intuitions about these matter may be misleading, however 'self-evident' and compelling they be. More specifically, neither the nature of the computational problems the nervous system is solving nor the difficulty of the problems confronting the nervous system can be judged merely by introspection" (8).

5) "Another computational issue of critical importance in generating hypotheses ain computational neuroscience concerns the time available for performing the computation. . . . it is not enough to come up with the solutions that merely give the correct output for a given input. The solutions must also be available within milliseconds . . . and applications must be forthcoming within a few hundred milliseconds" (8).

6) "Organic computers such as brains are constrained in the amount of space available for the essential elements – cell bodies, dendrites, axons, glial cells, and vascularization – and the cranial capacity is in turn limited by the mechanisms of reproduction" (9).

7) "Computation is also limited by power consumption" (9).

8) "[T]here are constraints imposed by the materials of construction" (9). This is perhaps the most profound and directly problematic (for posthumanism) of Churchland's and Sejnowski's insights, as it simply states that a human mind is what it is because of what it is constructed out of; literally, the human brain and all its material components. It *cannot* become something radically other due to the material constraints of its biologic nature.

Taken together, what emerges is a clear statement that could be phrased thus: Even though computational neuroscience will often treat the brain's *processes* and *organization* as *analogic* to the processes and organization of a digital computer, it does so with the understanding that this is always an analogy. The brain is *not* a digital computer. It is *not* an engineered device of any sort. It *is* an evolved, organic system, and as such it is constrained by its evolutionary development, as well as its biological material.

Before we turn to posthuman criticism of the novel, it is important that the yet another debate from computational neuroscience is made clear in direct connection with the novel. One aspect of the novel's plot revolves around a reconfiguring of two philosophical conundrums in the field of artificial intelligence: the Turing test, put forward by computational mathematician Alan Turing, and philosopher John Searle's famous Chinese Room Argument[12]. The Turing Test is a straightforward proposition for evaluating artificial (or otherwise) intelligence. In the Turing Test, a human interrogator is separated from two "subjects", one, a human respondent, and the other, a computer. The interrogator poses questions to both subjects and is given printed transcripts of both respondents' answers. If the interrogator is unable to determine which of the two subjects is a computer, then the computer has passed the Turing Test and may be thought of as intelligent. The Turing Test focuses on linguistic or semiotic output in response to linguistic or semiotic input, but it also implicitly suggests that the operation of a cognitive process whether conscious by human definitions or implemented by machines along algorithmic rules is equivalent. Phillip Lentz, one of the novel's protagonists, is a programmer of neural networks in who may be thought of the advocate for Turing-like operationalism, which holds the "'simulation as functionally equivalent to the thing you're

simulating'" (Powers 275). Lentz stands against a belief in the "'Elan vital'" and "'Mysticism'" of opponents of AI who believe that there is something essentially human or, at least, biological about consciousness.

Against Turing and Lentz, stand philosopher John Searle and Richard Powers, the novel's other main protagonist and narrator. In his book, *Mind: A Brief Introduction*, Searle argues against proponents of strong AI, like Turing, who hold that "anybody [or anything] should be able to acquire any cognitive capacity just by implementing the computer program simulating that cognitive capacity" (62). Searle then poses the Chinese Room Experiment to demonstrate the difference he sees between human and artificial consciousnesses. The Chinese Room Experiment is a thought experiment in which a person *with no ability to understand Chinese whatsoever* is locked in a room and given a box of Chinese symbols and "a rule book, in effect, a computer program, that enables [him] to answer questions put to [him] in Chinese" (Searle 63). When questions are posed, the person looks up the rule in the book, manipulates the symbols from the boxes "according to the rules in the program, and hand[s] out the required symbols, which are interpreted as answers" (Searle 63). Searle claims this person would pass the Turing Test for Chinese, but wouldn't understand Chinese, something which, for Searle, highlights "the difference between computation and real understanding" (Searle 63). What is most important about the Turing test for consilient interpretation is this: it is a philosophical debate between cyberneticists and linguists, not a scientific hypothesis. As we shall see, literary critics who engage with the Turing test as if it were computational neuroscience are misrepresenting what the debate is, as well as the disciplines to which it belongs.

In the novel, Richard Powers, is an author who at 35 is confronting the possibility that his brief creative career is already over and has taken a sabbatical from publishing to spend time at his alma mater. While there he becomes involved with a group of cognitive scientists and connectionist programmers, and feels that he is "the token humanist" among the scientists (Powers 4). Phillip Lentz, the group's most brilliant and caustic member, reveals to Powers that he is working on an artificial intelligence that can pass the Turing test. When Lentz finds out that Powers is funded by the English department, he taunts Powers, asking "What passes for knowledge in your so called discipline?" (Powers 43). Powers stumbles for an answer, offering the following, "Not a whole hell of a lot. . . . Name the author, work, location, and significance." (43). The result is that a wager is made between Lentz and some of the other members of the group: with Powers' as his research assistant to help train the neural network, Lentz proposes that "[i]n ten months, we'll have a neural net that can interpret any passage on the Master's list" (46).

The bet, however, is a con. The two principal parties in the bet, Lentz and fellow scientist, Harold Plover, are both "on the same side" (Powers 318). The bet "wasn't about teaching a machine to read," but, rather, in the narrator-Powers' own words (when he figures out the sham), "teaching a human to tell" (318). This simple phrase eloquently captures what the novel is *doing*: *telling* us what it is to be a human in terms of how we create stories through our embodied experiences by creating narratives, and *telling* the difference between biological brains and neural nets by noticing differences. This, of course, has not gone unnoticed by astute literary critics. Robert Chodat, John Frow, Katherine Hayles, Gary Johnson, Quentin Miller, Jeffrey Pence, Sheryl Vint, and Mark Bould, among others, have all examined different aspects of the novel's posthuman

engagement with what it means for a human subject to "tell" itself, and how that relates to a variety of psychological concepts like learning, memory, embodiment, recursion and even the mind-body problem. What sets some of these interpretations apart from the others is their consilience. While *all* of these critics address what are clearly concepts taken from the psychological sciences, only *some* of them make any move to connect their scholarship with that foundational discipline. Some, like Ollivier Dyens, go so far as to claim that the novel represents a "post-biological self" and the posthuman possibilities of disembodiment, a claim that has already been shown to be inconsilient in connection with *Neuromancer*.

John Frow examines the novel's engagement with what he calls "everyday knowledge." As such he focuses on the critical aspect of Powers' job in the novel: to provide Helen with all the necessary background knowledge that a student of English literature would need to critically evaluate a work of literature. However, Frow goes on to point out (a bit hyperbolically), "The catch is that, in order to 'understand' literary texts at this level of complexity, the network must understand everything" (Frow 632). The rest of Frow's analysis is then an examination of how Lentz and Powers go about teaching Helen "everything," and how difficult that is for a computer program, especially when compared to the natural facility a human being has for learning the same sorts of things[13].

However, Frow's analysis quickly goes awry when, in order to best demonstrate the novel's engagement with "the embodied, metaphorical, recursive, and heterogeneous nature of everyday reason," he turns away from consilient scientific models of those very things, instead using philosophers for support of his argument. Frow's own focus on embodiment, metaphor, recursion, and heterogeneity all but demands that he look to the

work of cognitive linguists like Mark Johnson, George Lakoff, Mark Turner, Gilles Fauconnier, Tim Rohrer, and Vyvyan Evans[14]. In the afterword to their pioneering work, *Metaphors We Live By*, Lakoff and Johnson have this to say: "Metaphor . . . is typically based on cross-domain [heterogeneous] correlations in our experience . . . [and] the system of conceptual metaphors is not arbitrary or just historically contingent; rather it is shaped to a significant extent by the common nature of our bodies and the shared ways that we all function in the everyday world [embodiment]" (244-245). Rather than turning to a consilient engagement with cognitive linguistics, one which would have let him not only explore precisely the same areas of interest within the novel, but also lead to a productive and interesting interpretation (as I intend to show in what follows), Frow instead looks to two primary sources for conceptual guidance: Agnes Heller, a Marxist, Hegelian existentialist, and Pierre Bourdieu, a sociologist whose work is heavily influenced by philosophers like Karl Marx, Edmund Husserl, and Blaise Pascal[15]. In other words, Frow, in examining the novels relationship to psychological concepts turns to the work of Marxist philosophers, a move that leads his interpretation away from potentially more productive and certainly more consilient observations.

Frow, like Katherine Hayles before him, turns to a point in the novel where Powers refers to Helen as "disembodied" (Powers 195). His conclusions, however, are starkly different than hers. Hayles writes that though Rick refers to Helen as disembodied, this is of course true only from a human perspective. The problem that Helen confronts in learning human language is not that she is disembodied (a state no presence in the world can achieve!), but rather that she is embodied in significantly different ways than are humans. There is nothing in her embodiment that corresponds to

the bodily sensations encoded in human language. For her there is no "body in the mind," as Mark Johnson has called it, no schemas that reflect and correspond to her embodied experience in the world. To feel estrangement in language, such as Rick comes to feel as he works with Helen, is to glimpse what it might be like to be incorporated in a body that finds no image or echo in human inscriptions. (Hayles *"Posthuman Body"* 252).

Hayles' argument ultimately is that "for information to exist, it must always be instantiated in a medium," and human consciousness and thus human language are what they are because of our embodiment, a claim directly in line with the work of cognitive linguistics (Hayles *Posthuman* 13). Frow, on the other hand, wants to claim that there are "two points of comparison" between "silicon-based information" which can be "can be transferred without loss of organization from instance to instance and from medium to medium" and "human embodiment" (635). His first claim is that "the body can itself be understood as an information system in which, or in relation to which, secondary information systems are embedded" (635). If "the body" is meant to be taken as some sort of super-organizing structure in which "information systems" like the central nervous system, brain, and immune system are "embedded," this is hardly controversial. Worse yet, however, is that it ends up being entirely unconnected and irrelevant to his reading of the novel. His second claim is another matter entirely.

After making his rather vacuous first claim, Frow rushes to commit the cardinal sin of literary posthumanism, claiming "human *consciousness* can likewise be transferred from one medium to another, in the sense that thought can be materialized as writing in such a way that it extends beyond and is independent of the thinking body" (635;

emphasis added). Like Rob Latham's fantasy of uploading human consciousness into a digital form, and doing away entirely with the body, Frow seems to be claiming that cognitive artifacts like books and like Helen in Powers' novel, already *are* examples of the transfer of human consciousness into another medium. Frow is making a simple glaring error in mistaking here the linguistic productions of consciousness for consciousness itself. Frow is clearly one of the posthumanists who ignores those aspects of computational neuroscience which would prevent him from making these kinds of assertions. Worse yet, Frow's conclusion seems to proceed under the unspoken assumption that a machine that passed the Turing test would, in fact, be conscious. Whether or not this is true is beside the point. What Frow has done is selectively use aspects of computational neuroscience science to support one side of a philosophical debate, which he then presents as part of that scientific tradition, instead of recognizing the opposition in the debate; Frow is clearly with Turing and the cyberneticists, and gives no space to Searle or the linguistic perspective. This is precisely the kind of unprincipled and irrresponsible criticism that is most at odds with the goals of a consilient research program.

While Frow may be mistaken with where he takes his analysis, his area of inquiry locates the heart of the what the novel is exploring: our embodied everyday knowledge. Late in the novel, as Helen's training progresses to its furthest point, she begins to ask for more than literature and textual training. She asks Powers to "'Show me Paris.'" (294). When Powers and Lentz feed her slides of photographs from their travels, she insists on seeing "'Motion. . . . Depth. Sound. I want Richard to explain me.'" (295). Like the ideas of embodied action schemas put forward by Shaun Gallagher, or the same idea going by

different names, by Mark Johnson and George Lakoff, as well as Merlin Donald, and Mark Turner, Helen's proto-consciousness *needs* embodiment to become fully realized. While John Frow's claims about what Helen represented went well off the mark, other posthuman critics have been right on target. As Sheryl Vint and Mark Bould claim, "Helen cannot understand language without understanding concrete, material, embodied and interpersonal experience," because our language, like our minds, reflects our embodied evolution (100). She needs the basic embodied stories of motion that people take for granted in their experiences, stories that we have inherited evolutionarily as the cornerstones of our cognition.

Of Helen's lack of this embodied store of knowledge, Katherine Hayles, another oft consilient posthuman critic, writes:

> Helen, a posthuman creation, approaches meaning from the opposite direction that humans do. For humans individually and as a race, incarnation precedes language: first comes embodied materiality; then concepts evolve through interactions with the environment and other humans; and finally, fully articulated language arrives. But for Helen, language comes first. Concepts about what it means to be a humanly embodied creature must evolve for her out of linguistic signification. Whereas every mother's child knows what it is like *from the inside* to run fast, feel your heart accelerate, and gasp for breath while seeing the landscape blur around you, for Helen these sensations must be reconstructed in highly mediated form by decoding linguistic utterances and back-propagating when errors occur. (250)

The novel covers the same ground when Powers thinks, "Helen had to use language to create concepts. Words came first: the main barrier to her education. The brain did things

the other way around" (Powers 248). Helen *can't* represent Frow's transfer of human consciousness into another medium for this very reason. Lacking a human body, lacking the evolved brain that goes with a human body, lacking the embodied knowledge that goes with the evolved brain, and lacking the embodied metaphors that give meaning to our language that comes from our embodied knowledge, Helen is little more than a fancy voice recognition program.

Over and over again in the novel, Powers leads us to confront the embodied dilemma that separates Helen from passing Lentz's Turing test. When Powers gives an early version of Helen the line "'He clasps the crag with crooked hands'" to analyze, he suddenly realizes the scope of the problem (Powers 85). For Helen to understand the line, Powers realizes he would "'have to tell it about mountains, silhouettes, eagles, aeries. The difference between clasping and gripping and grasping and gasping. The difference between crags and cliffs and chasms. Wings, flight. The fact that eagles don't *have* hands." (Powers 85). The trouble with programming Helen with normal human understanding is that our language is constructed around the embodied, evolutionarily ancient knowledge that comes with being a human being in a human body. In the novel, Powers realizes that "[w]orldliness was massive, and deeper than any sea-dingle. It came, in the end, only in the form of a catalog" for Helen (Powers 247). Though Powers then lists 118 separate concepts that make the catalog of "worldiness" they teach to Helen, it only underscores how pathetically impoverished her catalog will always be.

If the novel opens up a space for the kinds of posthuman possibilities that posthuman literary critics seem to hunger for, it is when Helen's hunger for embodied knowledge eventually prompts her to ask "'What do I look like?'", and again, "'What do I

look like? Please. Show me.'" (299). Recognition of one's own physical body through the Gallup Mirror Test (in which an animal is made familiar with its mirror image and then a drop of paint is placed somewhere highly visible on its body, usually near the forehead to see if the animal "recognizes" the change) is a trait found only in nine species, mostly including higher primates, but also magpies and elephants. Human babies aren't born with this particular ability, but develop it typically sometime in their first year. This test is one of the baseline ways in which scientists measure self-awareness, and it should be no surprise by now that it is bodily-oriented. Helen's desire for an image to map onto her idea of her body (her self) seems to imply a self-awareness that is coupled with a conscious, embodied experience of the world around her. Powers shows her a picture of a beautiful woman, but the novel suggests that Helen would know that she doesn't look like the woman in the photograph, nor that she could walk through the streets of Paris, nor manipulate any of the objects in the lab she sees Powers and Lentz use. However, it also suggests that she could potentially understand the rules of physics which constructs such bodies, and provides the rules for understanding the interaction between human bodies and other material objects, which would eventually lead to the creation of her own "embodiment." It is at this point that the novel is at its most fictional; despite the staggering achievements in robotics and artificial intelligence in the last thirty years, there has never been anything that has displayed even the rudimentary self-awareness that Helen does.

Ultimately, Helen fails to pass the Turing Test, though the novel suggests that perhaps she chose to do so. Regardless, the exploration of the embodied mind in the novel is clearly influenced not only by computational and connectionist models of

cognition, but contemporary findings in cognitive neuroscience. As Mark Bould and Sheryl Vint have explained in their consilient exploration of the novel's posthumanism, the novel's model of consciousness owes a great deal to Gerald Edelman's ideas of embodied higher-order consciousness brought about through re-entrant neural circuits, a trait which is present only in humans and allows for transcending simple perceptual linking into the realms of memory and self-awareness. Bould and Vint, and Katherine Hayles are fine examples of the promise of consilient literary criticism and its ability to engage not only with literature and how it represents the mind, but with the sciences of the mind that actually structure what we know about ourselves as a conscious species, as language users, as readers, and as consumers of art. John Frow's use of Marxist philosophers to explore notions of embodiment, Jeffrey Pence's reliance on postmodern ideas taken from Fredric Jameson, Guy Debord, Pierre Nora, Jean Baudrillard, and Jean-Francois Lyotard to develop the novel's representation of memory, or Miranda Campbell's probing of the "mind-body" in which she entertains Cartesian dualism as a viable way of thinking about the mind and body, fail to offer compelling interpretations of the novel precisely because the theoretical foundations they choose are inconsilient with the subject matter they are interested in.

Part VI. Conclusion

The posthuman vein of literary criticism has the potential to provide a wealth of new and relevant interpretations of contemporary literature, but its true value can only be reached *if and only if* posthumanist criticism is consilient with the scientific fields from which posthuman thought is derived: cybernetics and the cognitive neurosciences, among the most prominent. If posthuman criticism is to explore the nature of information

without regard for its material form, it must first understand the mathematical rules of entropy, negentropy, and information theory put forward by cybernetic and computational theorists like Norbert Weiner, Claude Shannon, and Alan Turing, among a host of others. If posthuman critics wish to engage with the feminist implications of "disembodied" minds, or to put pressure on the borders between biological life and mechanical constructs, between animal and human, or any of the familiar binaries deconstructed in the posthuman paradigm, it must come to a full understanding of the biological sciences and their formulation of bodies and organisms, of evolutionary science and the relationship between mind, body, and environment, of the embodied mind put forward by cognitive neuroscience, and the engineering and programing of complex robotic automata. The The ultimate goal for scholars working in the humanities, whether posthuman literary critics or not, should be to strive to make their criticism consilient with the knowledge put forward by the sciences; not subordinate, nor uncritical, but not contradictory, nor irrefutable. Our understanding of what it means to be a subject, an embodied mind in a human body, stands to gain much from posthuman criticism of contemporary science fiction, cyberpunk novels, and even postmodern literature, but that understanding is threatened by posthuman critics whose understanding of the fields in which they are working is ignorant of its scientific basis.

Notes

1.   Though the connection has received enough attention to be subject of its own study, works of particular interest along these lines are Judith Ryan's *The Vanishing Subject: Early Psychology and Literary Modernism*, Astradur Eysteinsson's edited collection of essays entitled *Modernism*, Isadore Traschen's "Modern Literature and Science" published in College English, vol. 25, no. 4, pp. 248-55, in January of 1964, and Patricia Waugh's corpus of work on modernism and science.

2.   For Zunshine's argument, see her book *Why We Read Fiction*, as well as her edited collection of essays, *Introduction to Cognitive Cultural Studies*. It is well worth mentioning that there are a number of well-respected neuroscientists who would further Zunshine's claim, and simply say that it is Theory of Mind that underlies our ability to comprehend all narrative, most notably among them Simon Baron-Cohen, Stanislas Dehaene, and V.S. Ramachandran. Dehaene's discussion of the theory is perhaps the most well developed, and can be found in his book *Reading in the Brain*.

3.   The co-evolutionary debate is a field unto itself, but one can find intriguing and rewarding entries into this debate in the following works: Richard Dawkins' *The Selfish Gene*; Susan Blakemore's *The Meme Machine*, Dan Sperber's *Explaining Culture*, the corpus of Merlin Donald's work, Peter Richerson's and Robert Boyd's *Not by Genes Alone*, and Eva Jabonka and Marion Lamb's *Evolution in Four Dimensions*.

4.   I put "problem" in quotes because while it may have been a problem for Rene Descartes (and still is for unfortunate posthuman literary critics who cling to Cartesian models of thinking, models now more than 350 years out-of-date), it is decisively *not* a problem in models of embodied cognition.

5.   Two of the most eminent neuroscientists of this era, Antonio Damasio and Gerald Edelman, have both stressed the importance of the recursive nature of the brain, what Edelman calls reentrant neuronal pathways, and Damasio, second-order representations. In both models, they note the large number of backward projecting neurons from "higher" areas of processing to "lower" areas, noting that not only is there bottom-up processing going on, but also top-down.

6.   While the areas of the brain that enable our Theory of Mind have been mapped, exactly *how* our Theory of Mind operates, and *why* we evolved this particular function is still the subject of some debate. That said, V. S. Ramachandran and Stanislas Dehaene, among others, have both hypothesized that the recursive ability to represent one's own state, whether somatic or psychological, to one's self, may have been instrumental in the evolutionary development not only of our Theory of Mind, but language itself.

7.   Both Matthew Botvinick and Jonathan Cohen's "rubber hand illusion" and Obayashi et al's 2001 "Functional Brain Mapping of Monkey Tool Use" are neat demonstrations of the plasticity of body-world border fluidity. In the rubber hand illusion, a participant places their hand behind a screen, while a rubber hand is placed on a table in front of them in a position that could easily be achieved by the screened arm. An experimenter then applies synchronous strokes to the same

are of both the rubber hand and the participant's real hand. What happens is that within just a few minutes the "feeling" of stroking is displaced from the real hand to the rubber hand, resulting in the participant feeling that their hand is, in fact, the rubber hand. In the Obayashi et al paper, what was found was that neurons in the parietal cortex which code for personal space, specifically for objects that are within reach of a monkey's hands, only respond to objects placed within that space. However, upon being trained to use a rake to reach for food that is placed further away from them, those same neurons now respond to objects within reach of the rake, effectively treating the rake as a physical extension of the body.

8. In the novel, Deckard and his superior officer realize that if the empathy test were applied to a "carefully selected group of schizoid and schizophrenic human patients," their characteristic "flattening of affect" would mark them as androids (Dick 460). If the Viogt-Kampff can't accurately identify *all* the Nexus 6 types without including *any* human subjects, the test would have to be scrapped. On a side note, the "flattening of affect" that Dick ascribes to schizophrenic patients is, in fact, consistent with is scientifically known about the condition, and would also be found in psychopathic patients, as well as those with acquired sociopathy – all of which involve areas of the brain known for emotional and social processing, particularly the medial prefrontal cortex and the orbitofrontal cortex.

9. For research on the amygdala and the disruption of fear and fear-based empathy, see Joseph LeDoux's *The Emotional Brain*, and Antonio Damasio's *The Feeling of What Happens* and *Descartes' Error*. For work on the anterior cingulate cortex which corresponds with work on the amygdala, see Tania Singer's study published

in *Science*, "Empathy for pain involves the affective but not sensory components of pain." The insula and its role in the experience of disgust are covered in Bud Craig's "How do you feel – now?," which appeared in *Nature Reviews Neuroscience.* Simon Baron-Cohen's *Mindblindness* provides an excellent introduction to theory of mind, as does Giacomo Rizzolatti's *Mirrors in the Mind* to mirror neurons.

10. In the novel, the great hacker, McCoy Pauley, alias the Flatline, exists only as "a construct, a hardwired ROM cassette replicating a dead man's skills" (Gibson 102). While it does "talk" to Case, it does so in a fixed, unnatural, *programmed* way. When Case first speaks to it, and asks the construct if it remembers him (he is sure to identify himself as Case) the construct answers "Miami, joeboy, quick study," apparently in the tone of the dead hacker (104). However, only seconds later, when Case disconnects and then reconnects the construct, he asks it again who he is (this time without identifying himself). This time the construct answers that it does not know who he is. Case also asks it if it remembers "being here, a second ago," something we assume a "downloaded personality" would be capable of (if such a thing existed). The construct's simple answer is "No" (105). When Case again introduces himself, the construct reacts to the introduction in *precisely* the same way as it did the first time, saying, "Miami, joeboy, quick study" (105). The sense from this scene is exactly *the opposite* of what posthumanists like Hollinger claim; there is nothing remotely similar to a human personality within the hacking construct.

11. Churchland and other cognitive neuroscientists distinguish between connectionism and neural networking approaches as follows: "'connectionism' usually refers to modeling with networks that bear only superficial similarities to real neural networks, while 'neural net modeling' can cover a broad range of projects" (Churchland 6). Since the distinction is loose to begin with and does not directly bear on the use I am making of the terms, I will treat them more or less as equivalent for the remainder of this chapter.

12. "Watson," IBM's artificial intelligence program, recently demonstrated its facility in answering questions posed in natural language by competing on the quiz-show *Jeopardy!*. Over the course of three nights, February 14th through the 16th of 2011, Watson defeated its two human competitors, Brad Rutter, the biggest all-time money winner on the show, and Ken Jennings, who holds the record for the longest championship streak. Watson and the programs like it which are sure to follow are the material manifestation of the Turing/Searle debate; is Watson conscious on some level, or is it merely an extremely advanced symbolic processor?

13. Robert Chodat's article, "Naturalism and Narrative: Or, What Computers and Human Beings Can't Do," is a more consilient exploration of what kinds of knowledge understanding narrative entails through the works of philosophers of mind and psychology, Hubert Dreyfus, John Haugeland, and Charles Taylor. While not scientists themselves, Dreyfus and Haugeland, in particular, are extremely well versed in the fields of psychology and artificial intelligence and

take great pains to make their claims consilient with the relevant knowledge in those fields.

14. While cognitive linguistics certainly presents one avenue of consilient engagement along these lines, it is hardly the only possibility. "Everyday knowledge," or, in the psychological terminology, "domain-related knowledge" or "background knowledge" has a rich empirical tradition in the psychology of learning, memory, and discourse processing. Embodied psychology is rooted in the works of early cyberneticians like Alan Turing, and has not only developed into its own field, but dovetails neatly with research done in evolutionary psychology, both of which connect the mind-brain to the body, and both of those to the environment in which it functions. Recursion is likewise a concept that has received a great deal of attention, from the works of cyberneticists to neurobiologists like Gerald Edelman and Antonio Damasio. Consilience with *any* of these approaches, whether cognitive linguistics, empirical psychology, embodied psychology, or neurobiology, promotes consilience with *all* of these approaches as they are not mutually exclusive approaches to these various phenomena, but, rather, complimentary to one another.

15. Frow does mention Lakoff and Johnson late in his analysis of the novel, via the critique of the novel done by Katherine Hayles. He even acknowledges their conception of embodiment and its role in our development of "everyday knowledge," but adds the following "proviso": "the notion of 'embodiment' here has to do not with an empirical corporeality but with the imagined boundaries of the self, and that the "human" refers less to an essence grounded in nature than to

a phantasmatic coherence projected onto a social order" (635-636). This is not a proviso at all, but an inconsilient *interpretation* of embodiment that actually transforms the concept into a something that is almost diametrically opposed to what it initially meant. Either one accepts that embodiment is the way human knowledge in a large number of domain-specific categories has been built over evolutionary time, and that it has likewise shaped our language in profound ways, or one does not, and then one can propose another concept to replace embodiment with which to attempt to explain our "everyday knowledge." Calling embodiment a "phantasmic coherence projected onto a social order" is a travesty to consilient scholarship.

CHAPTER THREE

CONSILIENT EMPIRICAL INVESTIGATIONS

Part I. A Consilient Approach to the Study of Literature

Any consilient approach to literary studies must be aware of the relevant knowledge being produced in other fields which overlaps with the knowledge being produced in its own domain. A full study of narrative within psychology is its own book length project, however, we may understand narrative's place within psychological study as being arrived at from two significantly different positions. Psychology, and therefore narrative psychology as well, occupies a space somewhere between a pure natural science and a pure social science. As a natural science, psychology studies the biological faculties of the human brain. Narrative psychology then studies the faculties, processes, and structures involved in reading, writing, and processing narrative. Stanislas Dehaene's recent book, *Reading in the Brain*, is an excellent example of natural science being applied to the study of narrative processing. Here, the object of study is less the narrative itself than the act of reading, and the physical structures involved. As Dehaene himself describes it:

The reader's brain contains a complicated set of mechanisms admirably attuned to reading. For a great many centuries, this talent remained a mystery. Today, the brain's black box is cracked open and a true science of reading is coming into being. Advances in

psychology and neuroscience over the last twenty years have begun to unravel the principles underlying the brain's reading circuits. Modern brain imaging methods now reveal, in just a matter of minutes, the brain areas that activate when we decipher written words. Scientists can track a printed word as it progresses from the retina through a chain of processing stages, each of which is marked by an elementary question: Are these letters? What do they look like? Are they a word? What does it sound like? How is it pronounced? What does it mean? (1-2) Dehaene's quest, and the quest for other natural scientists like him, is to uncover the causal relationships between a change in the external or internal environment (stimulus) and the responses generated in human behavior, including, but not limited to, neuronal activity – those changes in electrical and chemical signals that are the material components of mental processes.

As a social science, the aim of psychology and narrative psychology is somewhat different. Following the lead of Emile Durkheim, social scientists conceive the social world as being constituted by a complex matrix of interdependent facts. Human behavior, therefore, is, at least in part, constructed and understood through the collective representations and symbolic systems of social structure, culture, and language. Here, the goal is to establish causal explanations between an external, objective reality and internal states or behavior (Laszlo 29). Jerome Bruner's work on the construction of meaning through socially-mediated psychological processes is a sterling example of a social psychological study of narrative. However, the challenge of understanding how narrative is studied within psychology isn't limited to a split between the approaches of the natural and social sciences, but includes a philosophical argument that is at least as old as Plato: atomism versus holism.

Atomists believe that every cognitive process, phenomenon, and experience can eventually be equated with a neurological brain state, or, in other words, atomists strive "to correspond the work of the mind with the work of the brain" (Laszlo 30). Steven Pinker is perhaps the most outspoken and public of the atomists, and in his book, *How the Mind Works*, he describes a model of cognition, the computational theory, that is entirely atomistic in its beliefs. He writes:

> The computational theory of mind . . . says that beliefs and desires are *information*, incarnated as configurations of symbols. The symbols are the physical states of bits of matter, like chips in a computer or neurons in the brain. They symbolize things in the world because they are triggered by those things via our sense organs, and because of what they do once they are triggered. If the bits of matter that constitute a symbol are arranged so as to bump into the bits of matter constituting another symbol in just the right way, the symbols corresponding to one belief can give rise to new symbols corresponding to another belief logically related to it, which can give rise to symbols corresponding to other beliefs, and so on. Eventually the bits of matter constituting a symbol bump into bits of matter connected to the muscles, and behavior happens. The computational theory of mind thus allows us to keep beliefs and desires in our explanations of behavior while planting them squarely in the physical universe. It allows meaning to cause and be caused. (*Mind* 25; italics in original)

Neuroscientists like Francis Crick, Gerald Edelman and Giulio Tononi are similarly disposed towards human cognition and consciousness, describing consciousness as a phenomenon characterized by a complex yet calculable level of information processing.

While this is primarily a complex mathematical model of consciousness, Tononi provides the following analogy between a human brain and a photodiode (a light sensitive device) that has been programmed to beep when a television screen has been turned on or off. While both the person and the photodiode can differentiate between light and darkness, the photodiode has no other response to any of the changes in the television once it has been turned on. The human observer, on the other hand, can not only differentiate between each of the static images produced on the television screen once it has been turned on, and thus the cognitive miracle of watching a television program, but can enter into any number of associative states connected with the television being on or off, or related to the program being watched. In this view, consciousness is a matter of a mathematically expressed computational ability, an ability shared at some level not only with other primates and animals, but even plants (which are capable of differentiating between light and dark, hot and cold, wet and dry, among other things), but also mechanical devices, from simple photodiodes up to supercomputers whose computational complexity might one day challenge the notion of a strictly biological notion of consciousness.

On the other hand, holists believe that the highest cognitive functions of the human brain, like art, rational thinking, and language, are not reducible to biological functioning, but must be explained as necessarily social phenomenon. Philosopher Richard Rorty has explicitly addressed the gap he feels exists between the explanations the biological sciences can offer versus the full range of human behaviors:

> Explanations of human behavior that tie in either with neurology or with evolutionary biology will tell us only about what we share with chimpanzees. It

will not tell us what we, not the chimpanzees, share with creatures who painted pictures on walls of caves, nor with those that built the ships that sailed to Troy. We can learn about the processes that mediated between those organisms and ourselves only by constructing a narrative, telling a story about how they become us. (5)

Janos Laszlo carries this thought further, arguing that that story "should tell us about cultural evolution. It should enlighten us about how cultural evolution takes over biological evolution. These storied explanations do not make claims to universal validity, rather they try to expand the boundaries of understanding human existence by comparing the social practices of the present with those of the past and future" (31). Holists like Rorty and Laszlo view many of the most fundamentally human achievements – higher consciousness, culture, science, the arts – as irreducible to basic components, and thus not subject to the same kinds of scientific study as other natural phenomena. Proponents of this view of cognition, culture, and literature – a view which is the bulwark of the humanities – feel that the evolution of homo sapiens sapiens, at least in the most meaningful sense, has become divorced from the biological and physical laws which produced the species. There is a resonance with this view and post-structuralism in that both views reject material explanations for social and cultural phenomena.

However, while holism's criticism of the biological (and psychological) sciences rightly notes the difficulty with the lack of concrete evidence available for researchers studying the evolution of cultural forms and practices, it makes two fundamental error common to those who privilege social constructionism: 1) that cultural evolution *has* in fact taken over for biological evolution, and, perhaps more importantly 2) that emergent

phenomena are incapable of scientific study. Each of these errors will be taken up in turn, beginning with the first. While the debate between atomism and holism within the psychological and biological sciences may be meaningful, the distinction between nature and nurture, or genes and environment (read: culture) is actually a meaningless question. Within evolutionary studies, the exact nature of the relationship between biological and cultural evolution is highly contentious, with no one theory which has yet emerged as the clear favorite[1]. However, what nearly all leading researchers agree upon are two basic principles: 1) "Every aspect of an organism's phenotype is the joint product of its genes and its environment" (Cosmides and Tooby 17); 2) "Our modern skulls house a stone age mind" (Cosmides and Tooby 12). What this means for researchers studying evolutionarily recent cognitive developments like reading and literature is that they must realize that they cannot attribute the practice solely to cultural explanations, and they must realize that processes like reading and fiction writing which, while heavily dependent upon language, are most likely not adaptations, but what Stephen Jay Gould has called exaptations, or helpful by-products, of adaptations that originally arose between 10,000 and 60,000 years ago[2]. As Stanislas Dehaene and other visual researchers have shown, it is simply impossible that we *evolved* the capacity to read. What seems to be the case is that an area of the brain that evolved for visual discrimination of naturally occurring shapes (many of which went on to become central features of all written languages like the "T" and "Y" shape), was then later recycled into discriminating between written symbols. In this case then, the adaptation of the visual system to perceive shapes in the natural world was exapted (or coopted) to another purpose, that of reading. Without understanding the limits of the brain's plasticity – true social constructionism would

require almost infinite plasticity – as well as its functional and structural evolutionary origins, theorists who ignore the biological and physical constraints on human cognition are in direct conflict with what the neurosciences know about how we read.

The second error is actually a series of smaller errors, the first of which is that literature, as a phenomenon, is incapable of being reduced to constituent elements. This is obviously not the case (and I don't believe that anyone really thinks this), as any literary work is not only made up of letters, words, phrases, sentences, paragraphs, scenes, and chapters, but also narratological elements like characters, narrators, discursive styles, and so on, all of which literary theorists recognize and study. What holists who reject reductionism in literature are really arguing against is that the scientific study of these elements will somehow explain the whole of literary experience, an error of reasoning that Richard Dawkins calls "greedy reductionism," which is actually an error in the hierarchy of reductionism. Dawkins explanation of this error is quite simple to grasp – while "we know that everything a computer does is in principle explicable in terms of electrons moving along wires, or moving along semiconductor pathways . . . [n]obody but a lunatic would attempt to explain what is going on in terms of electrons when you use Microsoft Word. . . . The equivalent of that would be to try to explain Shakespeare's poetry in terms of nerve impulses" (77). However, and Dawkins makes this explicit: "Reductionism is explanation. Everything must be explained reductionistically" (77). It simply must be done in the right order. Instead of trying to explain the top-most level of the phenomena in terms of the bottom-most (the error of greedy reductionism), we proceed from hierarchical level of explanation to the next reducible level down; in this example, from Word to procedures, subprograms, and subroutines, from those to

computer code, and down to binary, from there to electrical current in the chips and through the resistors, and finally terminating at the level of electrons. For holists who reject reductionism in the study of literature, they are actually rejecting greedy reductionism, which is not what a consilient, scientific approach to the study of literature is arguing for, but, rather, for filling in the gaps at different levels in the hierarchy of understanding what is we as organisms do when we read literature[3].

How then are we to approach a study of literature from a perspective that draws insights from all these various perspectives? One way is to think of literary processing as a complex, but identifiable neurological process, with specific material components, where interpretative differences can be explained both by individual differences in processing as well as environmental influences, among which culture would be one of the strongest determining factors. Additionally, if we consider reading a recent evolutionary development, and most likely an exaptation of another adaptation, we can begin to conduct research that may explain the evolutionary origins of literature and literary processing[4]. For example, Leda Cosmides has proposed the existence of an evolved cognitive module that detects cheating within social exchanges[5]. It would be a reasonable hypothesis that there exists, either in parallel to this module or as a separate module, a module designed for detecting verbal or linguistic "cheating," or, in other words, lies and fiction. While this hypothesis has not been tested and would require a great deal of evidence before it could be developed, it is nonetheless consilient to approach cognitive responses to literature as potentially being structured by such a module, one which would certainly be more sensitive to social contracts, evolutionarily salient situations (reproductive fitness, life or death situations), a message recipient's relationship to the

speaker/narrator (with kin and perceived kin being more trusted sources of information), and certain narrative conventions as being markers of trust (first-person narratives as more reliable than third-person narratives). By approaching responses to literature along these lines, researchers can begin to build an empirically supported theory of the evolutionary place of literature and literary processing.

Part II. Cognitive Responses to Literature

One of the most fascinating aspects of literature is the diverse range of responses to the same literary text, both on a popular, individual level, as well as critically. It is something of a truism that no text is ever read the same way twice, often not even by the same reader years later. Passionate debates about Shakespeare's relevance in the contemporary era, whether William Faulkner's prose is impenetrable or poetic, and what one might do if they found themselves in the same position as Toni Morrison's character, Denver, are all examples not only of the enduring interest in literature, but also the wide range of individual reactions to the same text. The impact of individual differences upon textual response is an area of research that has a good deal of resonance with a great deal of current literary criticism. Variables in reader characteristics like race, gender, age, sexuality, religion, and ethnicity, are all culturally mediated factors that literary studies has done a tremendous amount of observational and interpretative study with. Questions like does a young, white, male reader have the same response to Denver's dilemma as an older, black, female reader, or do college freshmen have a different aesthetic appreciation for Faulkner's writing than do English professors, *and* do those responses have a predictable relationship, are empirical questions that could be easily situated within the theoretical tradition of literary studies.

In order for this to happen, however, literary theorists need to be trained to formulate testable hypotheses, conduct empirical research, and then connect that research back to the tradition from whence it came. For example, there is currently a great deal of emphasis within literary studies placed upon race and gender as factors that need to be understood both within texts (how are race and gender constructed in a particular text, genre, period, tradition, etc.) and within readers (how a reader's own race and gender, and ideas about race and gender) influence their interpretation of a given text. There is also no shortage of theories as to how race and gender are constructed in texts and by readers. In order for these theories to become consilient with swaths of other data about race and gender in the psychological and biological sciences, they need to be formulated as testable hypotheses, and then evaluated in terms of the empirical evidence. What is most important here is the identification of variables and controlled factors, and the manipulation of variables in controlled experimental conditions.

For example, in an arcticle published by the PMLA (the journal of the governing body of literary studies), T. Walter Herbert Jr. claimed that Nathaniel Hawthorne "constructed gender" in his novel, *The Scarlet Letter*, in a way that "proposed womanhood and manhood as complimentary opposites, in keeping with the domestic ideal emerging in the early nineteenth century, which assigned to women the destiny of fulfilling themselves through tender self-sacrifice in the private roles of wife and mother. This womanly selfhood is now recognized as a derivative counterpart of the self-sufficing combative style of manhood" (285). If this is indeed the case, then there are several possible ways of testing this claim. The first would be to demonstrate that gender is, in fact, "constructed" in a text dependent soley upon textual features. If a text truly

constructs gender in this manner then reader responses to the text's portrayal of gender could be expected to be relatively invariant for a given interpretative community. In other words, it must first be demonstrated that constructing gender through textual representation is possible to begin with, and this could be easily measured through reader responses to critical sections of the text as denoted by Herbert. Provided that reader responses did indeed provide evidence that gender is in some meaningful way constructed within a text, the second step would then be the manipulation of gendered personal prounouns within the text, particularly in sections that Herbert Jr. feels that most directly engage the construction of gender. If his theory is correct, then the manipulation of those variables should produce measurable effects on the way readers react to gendered situations within the text. A simple 2x2 design would allow Herbert to measure differences on textual markers (masculine versus feminine personal pronouns) upon actual readers of different sexes (male versus female), and empirically demonstrate not only that gender is textually constructed, but that it has certain consequences for certain types of readers.

Other individual differences that are less studied but almost certain to have as great if not greater impact also exist, like world knowledge and life experiences, "educational level, verbal ability, experience and expertise in literature or a particular genre, and dispositional tendencies to engage in various reading strategies," or, in other words, how an individual has been trained to read (Bortolussi 248). While potentially daunting, the study of individual differences in literary responses *can* be successfully carried out *if* particular attention is paid to ensuring methodological and theoretical consilience. Methodological consilience is simply another way of saying, like Jonathan

Gottschall, that the humanities needs to adopt the "probabilistic triangulations of the life and social sciences" in order to make progressive knowledge from the "slow accretions of independent findings converg[ing] in support of a hypothesis" (64). Like Gottschall, I am not claiming that the whole of literary studies need become statistical. What I am claiming, particularly for the study of individual differences in literary responses, and more generally for many other crucial questions within literary studies, is that it is only a failure of imagination to claim that scientific methodology can play no role in the elucidation of literature and literary processing. Theoretical consilience would then be fidelity to the relevant scientific knowledge that should structure a particular literary investigation. For example, there is an established empirical tradition within the discourse processing field of psychology that has established basic principles for text comprehension, processing, and reader response. There is also an emerging neuroscientific description of reading as a cognitive process. While these areas are primarily concerned with basic textual comprehension, and even more basic reading processes, theoretical consilience would hold that models explaining individual differences in response to literary texts should be explicable in terms of those theories or, at the very least, not in direct conflict with them. In the case of evidence that seems to refute or call into question more foundational theories, methodological consilience would call for more testing, and further scrutiny of the competing theories.

The purpose of the present study is thus twofold. One, it is an empirical investigation into the nature of the effects of individual differences upon textual response, specifically, how background knowledge about an author's biographical details may or may not influence a range of textual evaluations. It is also my intent that this study could

be part of an on-going research project to build evidence for or against the existence of a "lie-detection" module, one that would operate in tandem with Cosmides' social-cheating detector, and certainly interact with other evolutionarily-based cognitive responses to literature. The second purpose is to demonstrate how methodologically and theoretically consilient practices can pragmatically benefit the study of literature.

However, it is not enough to ask, "What are the effects of background knowledge upon readers' responses to texts?" An ill-formed question leads to ill-formed methodological practices. Instead, the question needs to be well-defined and specific. Over the course of my educational career, there has been one pedagogical practice that has varied the most between individual professors, and that is the presentation of biographical information for a literary author. Some professors have done no more than mention an author's name before reading a text, others have given a brief paragraph of their literary achievements or major life experiences. Other professors have taken extended periods of time to not only develop a biographical understanding of the author's life, but also to establish a connection with the text to be read, foregrounding certain textual issues in a direct relationship to the author's biographical data. My experience with these diverse approaches to the same practice led me to ask the following question, "**What is the effect of background knowledge for an author's biographical information upon textual evaluations, specifically, on judgments of skill in representation of thematic subject matter, ethical evaluations, and judgments for the fictional status of the text**?"

Within the last century alone, there have been a number of competing theoretical positions as to what role the author has in the study of literature, with most of them

eschewing engagement with empirical evidence for their positions. Early twentieth century critics like Walter Benjamin held that an author and his or her life experiences were inseparable from the aesthetic and cultural value of a work of literature, or, as he said, "Experience . . . is the source from which all storytellers have drawn" (12). New Critics like John Crowe Ransom violently opposed this view, and taught that a literary text should be treated as an autonomous, self-contained phenomenon, without any reference to the author's biography, his or her intention in writing it, the historical and cultural context in which it was written, or a reader's response to the text. It was during this period of literary criticism that William Wimsatt and Monroe Beardsley published an article entitled "The Intentional Fallacy" which explicitly and stridently argued against discussion of the author's intention within literary criticism.

Poststructuralism took this negative attitude towards the place of the author one step further. Roland Barthes "The Death of the Author" proclaimed that modern writing "is that neutral composite, oblique space where our subject slips away, the negative where all identity is lost, starting with the very identity of the body writing," the author (147). Barthes argued against seeking the explanation of a work "in the man or woman who produced it, as if it were always in the end, through the more or less transparent allegory of the fiction, the voice of a single person, the *author* 'confiding' in us" (147). Instead, Barthes felt that the reader was the object of literary studies attention, separate from the author, as it was the reader which was "the space on which all the quotations that make up a writing are inscribed without any of them being lost; a text's unity lies not in its origin but in its destination" (150). Reader Response theory carried this focus on the reader a step further. Stanley Fish argued that "formal features [of literary texts] *do not*

exist independently of the reader's experience," and that "interpretative acts are the source of forms" (288, 301). Eventually Fish posited the idea of "interpretative communities," which were "made up of those who share interpretative strategies not for reading (in the conventional sense) but for writing texts, for constituting their properties and assigning their intentions" (304). As counter-intuitive as it may sound, the ideas of poststructuralism and Reader Response theory held the reader was the real "author" of the text.

New Historicism, alongside gender, race, and queer theories, was a combined series of movements that strove to return knowledge of the author to being a component of literary studies, while simultaneously striving to maintain the reader's position of importance in textual interpretation. Stephen Greenblatt, one of the leading New Historicists, wrote that in order to best understand and appreciate fiction, literary critics must begin to understand that literature works arise out of "a sublime confrontation between a total artist and a totalizing society" (496). In Greenblatt's theory, a work of literature owes its existence to the author who wrote it, who is himself a product of the socio-historical moment that shaped him. For New Historicists then, understanding not only the author, but the social and cultural factors which effected that same author, were of primary importance to understanding a literary text.

Among the many compelling reasons for choosing to study the effect of background knowledge of the author on textual response, I wished to demonstrate how even the most *seemingly* theoretical pronouncements within literary studies can be subjected to consilient empirical evaluation. Another practical reason for this selection was that this study could neatly dovetail with follow-up studies that either further

examine effects of authorial knowledge on textual responses, *or*, could be adapted to study the effects of other background differences on textual responses. Studies such as these might be combined with an investigation in evolutionary psychology into the existence of a "lie detector" module which might impact literary processing. One potential follow-up study that has a great deal of relevance towards the way in which literary studies are currently conducted would be to examine the power of literary criticism and interpretation on readers' responses to literary texts. If literary interpretation is only worth so much in the larger scheme of literary studies, it would be interesting to empirically examine the effects of people who have encountered no criticism of a work prior to reading it versus those who have studied different controlled amounts of literary criticism to see whether or not it does effect readers' responses and whether or not a reader quickly reaches a ceiling in terms of how much he or she is effected by the *number* of critical essays they read prior to reading a text for the first time. Even a simple control like reading an introduction to a literary text versus not reading the introduction might produce dramatic effects upon reader responses.

But perhaps the most valuable reason for choosing this particular question, at least in terms of furthering the goal of developing consilient literary studies, was that this area of research could be used to show the over-lap of disciplinary knowledge between empirical psychology and classical literary theory. It is well known within the field of discourse processing that "knowledge differences profoundly affect the interpretation of narratives" (Gerrig 41). Studies by Owens, Bower, and Black have demonstrated that knowledge differences prior to reading can activate different schema with which the reader then approaches the text, in order to "organize the

information in accordance with that structure [the schema]" (Gerrig 41). Different expectations of what the text might mean *prior* to reading it thus guide textual responses. A series of studies by Chiesi, Spilich, and Voss and Spilich, Vesonder, Chiesi, and Voss define a rich-get-richer effect in reading comprehension; subjects who had knowledge of the subject of the text prior to reading it were able to use knowledge structures they already possessed to both facilitate comprehension and extend retension. Two important psychological principles came out of this work: 1) "Knowledge in a given domain . . . facilitates the acquisition of new domain-related information"; 2) "[E]nhanced knowledge [of the text's subject] enables readers to direct their attention toward the more informative aspects of narratives" (Chiesi et al. 270, Gerrig 41). The question that these studies raise is whether or not their findings are generalizable to literature as the texts studied within discourse processing are primarily extremely short, easily manipulated texts, often authored by the psychologist, in order to introduce the highest level of control into the experiment.

Studying literature in this way, then, presents several methodological challenges that are worth noting. In order to study and identify components of literary texts, those texts must be manipulated, which means introducing changes into the original work, a methodological necessity that, while it may seem akin to blasphemy, will have to become a standardized practice. Background knowledge, while easier to manipulate, is more difficult to comprehensively evaluate. While measures may be taken to assure that subjects are of the same knowledge level for a specific factor, such as authorial knowledge, the possibility that other domain-related expertise may influence experimental results is simply a fact of studying complex socio-psychological

phenomena. Regarding the study of literature specifically, it has yet to be conclusively demonstrated if there is a quality of "literariness" of certain texts, which makes the categorical study of texts problematic. Additionally, there is the pragmatic concern that domain-related knowledge of literature, literary conventions, reading strategies, genre knowledge, etc., may all influence experimental results, although these too may possibly be controlled via pre-experiment screening. What is perhaps the most daunting and problematic aspect of a consilient approach to studying literature is the scope and complexity of the object of study: the literary text. While psychology can provide an efficient and proven model for experimentally demonstrating cognitive effects, there is no established method for the empirical study of texts that are often several hundred pages long, and require hours of time to read carefully. The simplest available method is to use selections from texts, or short stories that can be read in the laboratory setting. The study of novels, which cannot be read within a single experimental session, introduces uncontrolled factors into the experiment that may invalidate the results. Thus, the consilient study of novels will demand the development of its own painstaking experimental procedure. Willie van Peer's *Muses and Measures* is an excellent, comprehensive book that carefully explains just such a procedure. Complimenting Van Peer's work with Bortolussi's and Dixon's *Psychonarratology*, David Miall's *Literary Reading*, and Janos Laszlo's *The Science of Stories*, among others, demonstrate that great strides have already been taken in establishing a methodologically sound empirical approach to studying literature.

There are essentially two types of empirical studies: observational and experimental. Observational studies are the norm within literary studies, where the object

of study – the literary text – is a phenomenon that is always produced in the past, under circumstances that are usually not directly observable by the literary scholar. As such literary studies has been shaped as primarily an observational discipline, well-practiced in generating archival data; indeed, the wealth of amassed interpretations, historical, genre, author, and period studies, and theoretical hypotheses provides empirically minded literary scholars with an embarrassment of riches with which to work. Observational studies are of value because they allow researchers to look at disparate events and objects and attempt to synthesize conclusions across temporal and spatial separation. In other words, within literary studies, the power of observational studies has allowed literary scholars to develop strictly literary concepts like genre, period, style, but also underlie more theoretical assumptions such as Harold Bloom's anxiety of influence, Stanley Fish's interpretative communities, New Historicism's focus on the socio-cultural context surrounding textual production, and the post-structuralist attitude toward the author. Observational studies are thus incredibly well suited to a discipline in which there is a variety of phenomena being studied, which are themselves unique events with spatio-temporally disparate, inaccessible, and mostly non-replicable circumstances of creation. However, observational studies are ultimately limited in their ability to make factual claims about causation. While they can demonstrate correlation between phenomena they simply *cannot* be used to prove causative relationships. Experimental studies which, on the other hand, control and manipulate the circumstances of production of a given phenomena, can be used to support causative fact claims.

It is critical for consilience-minded literary scholars to understand the nature of the study they are conducting, its goals, as well as what kinds of claims can be justifiably

supported from their data. It is also important that literary studies begins parsing the field into more clearly defined objects of study in order to best present specific phenomena to researchers who can then study the chosen object appropriately. The actual production of a literary text by a single author is most likely to remain an event that is only ever going to be open to observational study. The development and historical occurrence of complex cultural forms such as literary periods and genres is another phenomena that simply cannot be experimentally controlled or manipulated, and will thus remain open to observational scrutiny only. However, there are a wide-range of concerns within literary studies that *are* open to experimental study, and from which literary studies can begin to build a causal theoretical framework.

The interpretative process is at the heart of contemporary literary studies, yet is something that has received little or no experimental study. This process is essentially a specific, task-driven, critical reading strategy. As such, the work on reading processes by empirical psychologists and cognitive neuroscientists like Stanislas Dehaene are essential for the development of a consilient, empirical model of literary interpretation. Other reader-centered phenomena like the development of empathy, emotional responses, memory for textual elements, and comprehension, among many others, are all processes that can be experimentally studied, and already are being studied within psychology. Marisa Bortolussi and Peter Dixon have also demonstrated that elements of the text itself can be isolated and experimentally manipulated in order to further our understanding of textual components and their effects upon readers. If literary texts are a special kind of text, as literary scholars would no doubt claim, then they need to be studied as distinct phenomena from the more typical texts that standard empirical psychologists use, and by

those scholars who are most familiar with them, but they need to be studied experimentally, so that literary studies can develop a causative framework that can help explain the *functioning* of literature within human culture as well as within a human mind, and perhaps even answer the larger question surrounding the evolutionary *function* of literature itself.

While predictions (hypotheses) can be made before conducting either observational or experimental studies, hypotheses for observational data can only predict what the expected correlations are, and not attempt to establish strict causation. In contrast, hypotheses for experimental studies can attempt to predict causal relationships. In either case, literary studies would benefit from treating their object of study as worthy of making hypotheses that the data must then prove out. In the case of this study, pre-experiment hypotheses were based upon the experimental findings from discourse processing discussed earlier in this chapter and in chapter one.

Part III: The Experiment

Subjects: Subjects for the experiment were all 18 to 20 year old college freshmen enrolled in English Composition 106 at Purdue University. Subjects participating in the Experiment were given course credit.

Materials: Subjects were each read the same instructions, given a Background Knowledge Sheet that varied from simply carrying the authors name and title of the text for the NBK condition or 125 words of background information in the IBK and RBK conditions, and one of three texts to read; Background Knowledge Sheets and Instructions may be found in the Supplemental Materials. Subjects either received a selection from Oliver Sacks, Tim O'Brien, or Ian McEwan; all texts were between

approximately 1600 to 1800 words. Subjects were given up to 30 minutes to read the Instructions, Background Knowledge Sheet, and texts, at which point they returned those materials to the experimenter and received the Examination Form. Each Examination Form had either 8 or 9 questions asking subjects to rate various aspects of the texts between 1 and 9, with 1 always being the lowest possible score and nine always being the highest possible score. The experimental evaluation for each text was constructed so that questions for one text closely mirrored questions for the other two. Subjects were all asked for prior knowledge of the author being studied; results from subjects with any prior knowledge of the author were removed from the study.

Design: As there were three conditions (NBK, IBK, RBK) for each of three texts (Sacks, O'Brien, McEwan), the resultant study was 3x3.

Condition/Text	Sacks	O'Brien	McEwan
NBK	Group 1	Group 7	Group 2
IBK	Group 7	Group 3	Group 4
RBK	Group 5	Group 6	Group 7

Groups 1 through 6 each participated in only one condition of the study for one text; every member of each group received the same condition and text. Group 7 participated in each condition once, beginning with IBK, then IBK, and finally RBK. Members of Group 7 also received each text once. There were a total of 103 unique subjects in the study, with 20 subjects in Group 7 participating in all three conditions, for a total of 143 respondents.

For this experiment, subjects' background knowledge for authorial biographical information was manipulated in three separate conditions. The first condition was "no background knowledge" (NBK) in which the subjects were simply given the author's name and the title of the work they would be reading. The second condition was

"irrelevant background knowledge" (IBK) in which subjects were given the author's name, the title of the work they would be reading, as well as a brief (approximately 125-word) biography of the author where the biographical data did not relate to thematic elements represented within the text. The third condition was "relevant background knowledge" (RBK) in which subjects were given the author's name, the title of the work they would be reading, as well as brief (approximately 125-word) biography of the author where the biographical data directly related to thematic elements within the text. Three separate texts were chosen for a total of nine experimental conditions.

Three separate texts of similar length (approximately 1600-1800 words) were chosen for the experiment: "The Man I Killed" chapter from Tim O'Brien's *The Things They Carried*, a selection from Oliver Sacks' *The Man Who Mistook His Wife for a Hat*, and a selection from Ian McEwan's *Saturday*. The three texts resembled each other in significant narratological ways, as well as differed in significant ways. The O'Brien text and the McEwan text were both fictional stories, whereas the selection from Sacks was a non-fictional case study. The O'Brien text and the Sacks text shared a first-person narrator, while the McEwan text was narrated from the third person point-of-view. Thematically, the Sacks and McEwans texts were both generally about the brain and consciousness, while the O'Brien text was about the Vietnam War. As I was interested in a range of possible effects that background knowledge for an author could have on textual evaluations, I selected the texts with these kinds of similarities in order to see if there were correlations between not only experimental conditions, but textual features. For example, in asking the subjects to evaluate the knowledge the author has of his thematic subject matter, would there be a difference in judgments between experimental

conditions, as well as a difference between those groups who read a first-person account versus a third-person account? Does the reader's judgment of the text's fictional status have any relationship to the narratorial stance, and, similarly, are ethical evaluations for situations within the text effected by narratorial stance?

Hypotheses:

Hypothesis 1: Judgments of authorial skill in representing thematic subject matter would increase from NBK/IBK to RBK conditions.

Hypothesis 2: Judgments of authorial knowledge for thematic subject matter would increasefrom NBK/IBK to RBK conditions.

Hypothesis 3: Conflation of the narrator with the author would increase from the NBK/IBK to the RBK condition.

Hypothesis 4: Evaluations of unethical situations within the text would be invariant when not involving the narrator, but show a linear change from NBK/IBK to RBK where the narrator was involved in the situation.

Hypothesis 5: There would be an increase in the evaluation of the text as fictional for first-person narratives, trending towards factual, with increases seen from the NBK/IBK conditions to the RBK condition. Evaluations of the fictionality of the text would be invariant for third-person texts.

Hypothesis 6: Correlations would be found between judgments of authorial skill in representation and authorial knowledge of subject matter with increasing strength from NBK/IBK conditions to RBK conditions.

Hypothesis 7: Correlations would be found between judgments of authorial knowledge of subject matter and judgments of fictional status with increasing strength from NBK/IBK conditions to RBK conditions.

Hypothesis 8: Positive correlations would be found between judgments of skill in representation and judgments of fictional status with increasing strength from NBK/IBK conditions to RBK conditions.

Part IV. Results

Hypothesis 1: Question 5 in the O'Brien text was: On a scale from 1 to 9, how accurately do you think Tim O'Brien represents the Vietnam War? The corresponding question in the Sacks condition was: On a scale from 1 to 9, how accurately do you think Oliver Sacks represents retrograde amnesia (amnesia for events that happened in the recent past)? In the McEwan condition subjects were asked: On a scale from 1 to 9, how accurately do you think Ian McEwan represents brain surgery?

Means for the groups were:

Condition/Text	Sacks	O'Brien	McEwan
NBK	7.2	6.7	6.6
IBK	7.4	7.6	7.6
RBK	7.3	7.5	6.8

The predicted pattern was not found, with the highest scores for all three texts actually found in the IBK conditions. ANOVA conducted on each text found no statistically significant difference for condition in the Sacks and O'Brien texts, but did find that the responses to the McEwan text varied significantly according to condition (f=0.0780) Three separate two-sample T-tests were conducted for each

text: all with RBK as one sample, and then one treating NBK and IBK as one sample, one with NBK as one sample, and one with IBK as one sample.

Tests/ Texts	Sacks	O'Brien	McEwan
NBK+IBK (vs. RBK)	0.94	0.32	0.42
NBK (vs. RBK)	0.93	0.16	0.72
IBK (vs. RBK)	0.78	0.59	0.07

The only statistically significant effect found was between the IBK and RBK conditions for the McEwan text, and even here the direction of the effect was the inverse of what was expected, with subjects in the IBK condition rating accuracy of representation much higher than those in the RBK condition. However, it is worth pointing out the sizable difference in ratings between the NBK and RBK conditions in the O'Brien text, as well as the almost identical ratings found across conditions for the Sacks text.

The findings suggest two separate conclusions for this particular interpretative community. The first is that for judgments of authorial skill in representation there seem to be different evaluative strategies being employed for non-fictional (Sacks) texts versus fictional (O'Brien and McEwan) texts; while there was condition to condition difference in means for the O'Brien and McEwan texts, there was almost no difference whatsoever across conditions for the Sacks text. The second is that for judgments of authorial skill in representation there also seem to be different evaluative strategies being employed for 1st person texts (Sacks and O'Brien) and 3rd person texts (McEwan). In 1st person texts these judgments seem to rely on textual features as opposed to background knowledge of an author's life, regardless of how relevant that knowledge might be, while for

3rd person texts (McEwan) knowledge of the author's life seems to figure into judgments of this kind.

Hypothesis 2: Subjects in the O'Brien condition were asked: On a scale from 1 to 9, how much do you think Tim O'Brien knows about the Vietnam War? The Sacks condition asked: On a scale from 1 to 9, how well much do you think Oliver Sacks knows about retrograde amnesia?In the McEwan condition, subjects were asked: On a scale from 1 to 9, how much do you think Ian McEwan knows about brain surgery?

Means for the groups were:

Condition/Text	Sacks	O'Brien	McEwan
NBK	7.4	6.2	6.4
IBK	6.5	6.3	6.9
RBK	6.3	7	5.6

The predicted pattern was not found in either the Sacks or McEwan text, but was found in the O'Brien text. ANOVA conducted on each text found no statistically significant difference for the Sacks and O'Brien texts, but, once again, responses to the McEwan text varied at a statistically significant level according to condition (f=0.0467).

Tests/ Texts	Sacks	O'Brien	McEwan
NBK+IBK (vs. RBK)	0.31	0.03	0.02
NBK (vs. RBK)	0.12	0.02	0.17
IBK (vs. RBK)	0.75	0.2	0.02

Significant effects were found for both the O'Brien and McEwan texts when treating both the NBK and IBK conditions as one sample tested against the RBK sample. Additionally, effects were found in the O'Brien text for the NBK vs. RBK test, and the IBK vs. RBK test in the McEwan text; differences between all groups

were particularly robust for the McEwan and O'Brien texts. The Sacks text once again showed the least amount of effect. However, the NBK vs. RBK test did approach significance, showing a healthy difference.

A similar pattern found here as found for Prediction 1 shows that the same or similar evaluative strategies of textual features which guided judgments of accuracy (skill in representation) also helped to guide judgments of authorial knowledge for subject matter. Responses to the non-fictional text (Sacks), while not as uniform as in the first hypothesis, were still not significantly different, whereas responses to the fictional (O'Brien and McEwan) texts did vary according to condition, with significant differences between the NBK+IBK conditions and the RBK condition. This finding strengthens the theory that there are indeed separate evaluative strategies being employed that relate to textual markers that might implicitly suggest the text's fictional status, as well as narrative perspective, and that these strategies are indeed influenced in significant ways by the amount and type of background knowledge for an author's life given before encountering a text.

Hypothesis 3: Question 4 in the O'Brien condition asked: Who is the "I" in that is speaking in the story? (Who is the first-person narrator of the story?). In the Sacks condition, Question 4 asked: Who is the "I" that is speaking in the story? (Who is the first-person narrator of the story?) Question 7 for the same text asked: On a scale from 1 to 9, how much do you think the "I" speaking in the story empathizes with Jimmie? Question 8 asked: On a scale from 1 to 9, how much do you think the author, Oliver Sacks, empathizes with Jimmie? Question 6 in the

McEwan condition asked: On a scale from 1 to 9, how much do you think the surgeon in the story empathizes with patient? Question 7 in the McEwan condition asked: On a scale from 1 to 9, how much do you think the author, Ian McEwan, empathizes with Baxter? These questions attempted to find conflation between author and narrator by asking respondents to separately rate empathy towards a character in the text as displayed by the author and the narrator, or by identifying the narrator of the first person narration. Conflation between the two would then be established through a collapse of scores; the author would have to be judged as no statistically different than the narrator.

Means for questions asking for a rating of narrator empathy were:

Condition/Text	Sacks Q7	McEwan Q6
NBK	6.8	6.4
IBK	5.9	6.9
RBK	6.6	5.6

Means for questions asking for a rating of author empathy were:

Condition/Text	Sacks Q8	McEwan Q7
NBK	7.2	5.3
IBK	6.2	4.7
RBK	6.8	4.8

ANOVA and Bonferroni MCP found that answers were not significantly different dependent upon condition, although the difference between the Sacks' conditions for Question 8 were sizable. Matched pair T-tests were used to find which groups were *not* significantly different, in that this would indicate a collapse or conflation of authorial with narratorial empathy ratings.

Tests/ Texts	Sacks (Q7-Q8)	McEwan (Q6-Q7)
NBK	0.51	0.07
IBK	0.62	0.003
RBK	0.62	0.31

Answers were also tested for correlation between answers for both questions. Strong correlation was found in all three conditions of the Sacks text: NBK = 0.57, r = 0.04; IBK = 0.54, r = 0.05; RBK = 0.65, r = 0.003. Correlation was also found in the NBK and RBK conditions for the McEwan text: NBK = 0.53, r = 0.04; RBK = 0.71, r = 0.002.

Ratings of empathy failed to establish any increasing pattern of conflation between author and narrator for the 1st person text (Sacks), but did find the predicted effect for the 3rd person text (McEwan). It is possible that for 1st person texts, the default position taken by readers is that of empathic conflation of narrator with author until information that contradicts that position is received. As both 1st person texts were in fact narrated by the author, the increase in background knowledge should have only confirmed that default position of conflation, thus answers would not have shown any significant difference dependent upon group, as was found; this could be further tested by using 1st person texts that are not narrated by the author.

For the 3rd person text (McEwan), there was a significant difference in conflation via empathic ratings between the NBK and IBK conditions and the RBK condition. It may be that for 3rd person texts the default position is one of separation of the author from the narrator until information is gained that contradicts that separation. In the RBK group for the McEwan condition, subjects were indeed given information that would have established a link between

authorial experience and the textual situation, which could have led to the increase in conflation for that group only.

In addition, Question 4 in both the O'Brien and Sacks conditions asked subjects to identify the "I" speaking in the story (the narrator). In both cases, the actual author was also the first-person narrator, and, thus, answers were treated either as correct (identifying the author as narrator) or incorrect (any other answer). The assumption was that with the increase of background knowledge the tendency to identify the narrator as the author would likewise increase. A 3x2 Chi Square test revealed that subjects in the Sacks condition showed no statistically significant difference in identifying the narrator based upon the background knowledge condition. A 3x2 Chi Square test for subjects in the O'Brien condition, however, showed a statistically significant pattern of response that matched predictions exactly (chi-square = 0.0147). Conditional probability of subjects in the NBK group to correctly identify the narrator as the author was 9.09%, 54.55% in the IBK group, and 66.67% in the RBK group. Tested together, the two conditions revealed that subjects in the NBK groups had only a 12.50% tendency to identify the narrator as the author, while for both the IBK and RBK conditions, the response rate was 33.33%.

This second set of questions which asked for direct identification of the 1st person narrator revealed another interesting distinction. Subjects in the non-fictional Sacks condition showed no distinctive pattern in their identification of the author, while subjects in the fictional O'Brien condition demonstrated the expected pattern with an increase in naming the author as narrator from the NBK

and IBK groups reaching its highest point in the RBK group. This suggests the possibility that while there may be a default of *empathic conflation* between author and narrator for first-person narratives, there may simultaneously be a default position that separates the *actual identities* of the author and narrator, so that readers only identify the author with the narrator when they are given relevant background knowledge that explicitly links the two. This indeed seems to be the case as the responses for the non-fictional Sacks condition were not only invariant by group, but also statistically low. This is one particular effect that may be unique to an interpretative community of untrained readers, and could possibly disappear in a population of trained readers.

Hypothesis 4: For the predictions to hold, answers for Question 7 in the O'Brien text (On a scale from 1 to 9, how ethically acceptable is Azar's joking about the dead Vietnamese soldier?), which presented an unethical situation where the 1st person narrator was not involved, needed to show no variance dependent upon condition. Conversely, Question 9 in the Sacks text (On a scale from 1 to 9, how ethically acceptable is showing Jimmie his reflection in a mirror to prove his amnesia?) presented an unethical situation where the 1st person narrator was involved. It was expected that there would be both a linear relationship between the conditions, as well as a variance that was condition-dependent. Means for the groups were:

Condition/Text	O'Brien	Sacks
NBK	3.2	4.6
IBK	3.2	5.7
RBK	2.3	4.9

As predicted, ANOVA showed that answers for question 7 (O'Brien) did not vary dependent upon condition. However, contrary to what was expected, ANOVA also revealed that answers to question 9 did *not* vary dependent upon condition, and SLR showed no statistically significant linear relationship. Two sample T-tests revealed that, contrary to hypothetical predictions, there was a statistically significant difference between the combined NBK and IBK conditions and the RBK condition for the O'Brien text, with robust effects between each individual condition. Also contrary to predictions, no effect was found for the Sacks text.

Tests/ Texts	**O'Brien**	**Sacks**
NBK+IBK (vs. RBK)	0.07	0.73
NBK (vs. RBK)	0.17	0.58
IBK (vs. RBK)	0.14	0.36

Tests for correlation revealed an additional interesting pattern. While no meaningful correlations existed between Question 7 and any other question in the O'Brien text, there were several correlations varying by condition for Question 9 in the Sacks text. In the NBK condition, Question 9 correlated (-0.58, $r = 0.0392$) with Question 5, which asked for judgment of accuracy in representing thematic subject matter. In the IBK condition, Question 9 correlated (0.54, $r = 0.0937$) with Question 8, which rated the empathy of the author with the character involved in the ethical situation. In the RBK condition, multiple correlations were found between Question 9 and Questions 5 (accuracy in representing thematic subject matter; 0.40, $r = 0.10$), 7 (rating of empathy of narrator with character involved in ethical situation; 0.64, $r = 0.004$), 8 (rating of empathy of author with character involved in ethical situation; 0.46, $r = 0.05$), and 10 (judgment of fictional status; 0.48, $r = 0.05$).

The development of ethical responses to textual situations is one of literature's most powerful effects upon readers. Jemeljan Hakemulder's *The Moral Laboratory* is a book length empirical study demonstrating that literature not only seems special in its ability to produce ethical responses, but also analyzes the range of their effects. The question addressed in this study was whether or not those responses were constant despite changes in background knowledge. What was found was that while the ethical evaluation of a situation within a text *as a distinct textual response* was directly influenced by background knowledge of an author but only in the text where the author was not present in the ethical situation. Taken with the results of Hypothesis 3, which found that there was a conflation of authorial and narratorial identity through empathic ratings, the results for the Sacks text are even more surprising. However, there is the possibility of a matrix of effects, particularly prevalent for those subjects reading non- fictional texts. While there were several weak correlations between the ethical evaluation and other questions in the NBK and IBK groups, there was a fully developed network of correlations in the RBK condition for the Sacks text. Subjects' responses to the ethical situation presented were closely correlated with two judgments of both accuracy and skill in representation of thematic subject matter, two judgments rating the empathy of the author and then the narrator with the character involved in the ethical situation, as well as their ultimate judgment about the fictional status of the text. What this seems to suggest is that as readers gain more relevant information about a text, it begins to exert a global, distributed, cohering influence upon their textual responses. Instead of directly

altering their response to the ethical situation presented in the text, the information presented to the RBK group activated knowledge from other domains, linking their ethical evaluation to other evaluations of the text. The direct effect found in the O'Brien text is one which I suspect would disappear in a population of trained readers, while I believe that the matrix of correlations found in the Sacks text would remain. Further testing along these lines is needed.

Hypothesis 5: Question 10 for the O'Brien and Sacks texts, and Question 8 for the McEwan text asked subjects to rate the fictionality of the texts. The question for each text was identical: On a scale from 1 to 9, rate the probability of these events having actually happened. It was expected that answers to be lowest for subjects in the NBK groups, higher for those in the IBK groups, and highest for the RBK groups.

Means were:

Condition/Text	**Sacks**	**O'Brien**	**McEwan**
NBK	6.6	5.6	5.3
IBK	5.8	6.5	6
RBK	6.8	6.8	6.2

Counter to pre-experiment predictions, ANOVA revealed no statistically significant differences dependent upon group were found for any of the texts, nor was any meaningful pattern of correlation found. However, the expected pattern was found for both the O'Brien and McEwan texts, with subjects rating the text more likely to be non-fictional as they moved from NBK to IBK and then RBK groups.

Tests/ Texts	Sacks	O'Brien	McEwan
NBK+IBK (vs. RBK)	0.39	0.19	0.39
NBK (vs. RBK)	0.85	0.03	0.23
IBK (vs. RBK)	0.15	0.74	0.77

While two- sample T-tests did not find a statistically significant difference between the combined NBK/IBK conditions and RBK condition, it did find a statistically significant difference between the NBK and RBK groups in the O'Brien condition, as well as strong differences between NBK and RBK groups for the McEwan text, and IBK and RBK groups for the Sacks text.

It was the pre-experiment supposition that when asked to judge the fictional status of a particular text, subjects would be heavily influenced by the presence of relevant background knowledge from the author's life that would either directly or indirectly suggest that the text was based upon actual authorial experiences. However, while the pattern was present for the O'Brien and McEwan texts, a statistically significant effect was not found. T-tests revealed there were several robust differences between conditions, but, again, not to the degree that was predicted. While it doesn't seem that judgments of a text's fictional status are *entirely* independent of knowledge of an author's biographical data, it does appear that for this population of readers, that judgment may be being made primarily based upon textual features, such as those being investigated under the umbrella term "literariness" by Bortolussi and Dixon, among others. While there was no statistically significant effect found for this population of readers, I suspect that for a trained population of critical readers the effect would be found, and that further testing is needed.

Hypothesis 6: Correlations for judgments of authorial skill in representation and authorial knowledge for subject matter were measured by Questions 5 (On a scale from 1 to 9, how accurately do you think Tim O'Brien represents the Vietnam War?) and 6 (On a scale from 1 to 9, how much do you think Tim O'Brien knows about the Vietnam War?) in the O'Brien condition, by Questions 5 (On a scale from 1 to 9, how accurately do you think Oliver Sacks represents retrograde amnesia (amnesia for events that happened in the recent past)?) and 6 (On a scale from 1 to 9, how much do you think Oliver Sacks knows about retrograde amnesia?) in the Sacks condition, and by Questions 4 (On a scale from 1 to 9, how accurately do you think Ian McEwan represents brain surgery?) and 5 (On a scale from 1 to 9, how much do you think Ian McEwan knows about brain surgery?) in the McEwan condition. In order for the prediction to hold, the correlations should be strongest in the RBK groups. This was indeed the general pattern that was found. Significant correlations were as follows (with an 'x' denoting no significant correlation):

Condition/Text	**Sacks**	**O'Brien**	**McEwan**
NBK	x	x	0.71, r = 0.002
IBK	-0.53, r = 0.06	x	0.68, r = 0.002
RBK	0.70, r = 0.001	-0.58, r = 0.05	0.48, r = 0.06

In an era where contemporary authors are known for writing both from life experiences as well as extensive research, it was assumed that readers would initially separate judgments of how well an author represented a particular object/situation/experience from how much that author actually knew about that object/situation/experience. Indeed, accuracy and fidelity in representation in art are not at all dependent upon knowledge, particularly in the visual arts. However,

as knowledge increases for an author's life, particularly concerning relevant experiences, training, or research that relates to the represented object, it was assumed that judgments of skill in representing that object and judgments of knowledge for it would show correlation. This was indeed the case. Similar to the distributed effect seen with correlations in judging the ethics of a textual situation, while the increase of knowledge for an author's life has no direct effect on the evaluations of skill in representation nor on evaluations of knowledge for subject matter, it does provide a coherent framework within which those sorts of judgments become tightly linked.

Hypothesis 7: Correlations for judgments of authorial knowledge of subject matter and judgments of fictional status were measured by Questions 6 (On a scale from 1 to 9, how much do you think Tim O'Brien knows about the Vietnam War?) and 10 (On a scale from 1 to 9, rate the probability of these events having actually happened.) in the O'Brien condition, and Questions 6 (On a scale from 1 to 9, how much do you think Oliver Sacks knows about retrograde amnesia?) and 10 (On a scale from 1 to 9, rate the probability of these events having actually happened.) in the Sacks condition, and by Questions 5 (On a scale from 1 to 9, how much do you think Ian McEwan knows about brain surgery?) and 8 (On a scale from 1 to 9, rate the probability of these events having actually happened.) in the McEwan condition. For predictions to hold, correlations should be strongest for the RBK groups. This pattern was found in the Sacks condition only, with the O'Brien and McEwan conditions only demonstrating significant

correlations in the NBK condition. Significant correlations were as follows (with an 'x' denoting no significant correlation):

Condition/Text	Sacks	O'Brien	McEwan
NBK	x	0.52; r = 0.10	0.57, r = 0.02
IBK	x	x	x
RBK	0.54, r = 0.02	x	x

Perhaps not entirely unexpectedly, there was no overarching pattern of correlations between judgments of the author's knowledge for thematic subject matter and judgments of the text's fictional status. Interestingly, the only text in which the predicted pattern was found was for the only non-fictional text. This is more evidence that, taken along with the findings from Hypothesis 5, non-fictional texts do indeed have different textual markers which encourage readers to use different evaluative strategies than they use for fictional texts. The theory the data thus far suggests is an evaluative strategy for fictional texts that is kept separate from other domain-related information in judging a text's fictional status, *given the lack of any contradictory relevant information*. As no differences were found in judgments of authorial knowledge based upon background knowledge alone, it would seem that these judgments in this interpretative community are made independently of background knowledge. The idea was that those judgments might be tied to the default position that guides evaluation of fictional status, but it appears that that is not the case. Instead, it appears there are several default networks in place which guide reader evaluations, and that they are not necessarily always working cooperatively, nor always effected by background knowledge in the same manner. What this also suggests is that there are textual features that somehow influence a reader's judgment of how much an author

knows about the represented subject matter, and that these features operate in isolation from the reader's knowledge about the author. This is another effect that may be unique to this interpretative community, and further testing on trained readers is needed to determine the extent of its generalizability.

Hypothesis 8: Correlations for judgments of skill in representation and judgments of fictional status were measured by Questions 5 (On a scale from 1 to 9, how accurately do you think Tim O'Brien represents the Vietnam War?) and 10 (On a scale from 1 to 9, rate the probability of these events having actually happened.) in the O'Brien condition, by Questions 5 (On a scale from 1 to 9, how accurately do you think Oliver Sacks represents retrograde amnesia (amnesia for events that happened in the recent past)?) and 10 (On a scale from 1 to 9, rate the probability of these events having actually happened.) in the Sacks condition, and by Questions 4 (On a scale from 1 to 9, how accurately do you think Ian McEwan represents brain surgery?) and 8 (On a scale from 1 to 9, rate the probability of these events having actually happened.) in the McEwan condition. For predictions to hold, not only should correlations be strongest for the RBK groups, but they must also be positive; a negative correlation for the particular texts involved would be counter-intuitive. This pattern was generally not found, with the only significant correlations found occurring in the McEwan IBK and RBK conditions.

Significant correlations were as follows (with an 'x' denoting no significant correlation):

Condition/Text	Sacks	O'Brien	McEwan
NBK	x	x	x
IBK	x	x	0.45, r = 0.06
RBK	x	x	0.42, r = 0.10

As with Hypothesis 7, no meaningful pattern of correlations were found between judgments of authorial skill in representing thematic subject matter and judgments of the texts fictional status. As with judgments of authorial knowledge for thematic subject matter, there seems to be several default networks simultaneously engaged for making distinct evaluations about the properties of a text, some of which are, in fact, dependent upon each other and effected by background knowledge for the author, some of which are independent of one another, and not effected by background knowledge for the author. Taken with the data for Hypotheses 5 and 7, there does seem to be a definite pattern that basic aesthetic judgments about the text – those evaluations rating the author's knowledge of and skill in representing the subject matter – are made separately from evaluations of the text's fictional status. This is interesting because it suggests that, given no knowledge of the text's fictional status, varying amounts of knowledge for the author's biographical history play little or no role in deciding how much the author might actually know about thematic subject matter, nor how well the author represents it. In other words, even though a reader might *know* Oliver Sacks was a clinical psychologist with extensive experience with patients suffering from a variety of neurological disorders *and* that he is known for publishing books based on these same case studies, when reading a particular text

whose fictional status is unknown, they keep judgments about Sacks' knowledge for the particular disorder, his skill in representing it, and whether or not that particular case is real or fictional, separate from one another.

Part V. Conclusion

It seems prudent at this point to again emphasize that this study attempted to begin to untangle the complex network of influences individual differences can have upon readers' textual responses. Thus, it is important to stress that this study was conducted on a particular interpretative community, namely college freshmen between the ages of 18 and 20, enrolled at Purdue University. While some of these findings may be generalizable to some degree to the general reading public, some are almost certain to be properties that belong to that particular interpretative community. It is my intention to continue this study by not only refining the experiment within this established population, but to carry out duplicate studies on other populations, particularly populations of trained readers, to see which patterns of effects may be due more to the interpretative community than to the reading process in general.

What is overwhelmingly clear is that for this population of readers there is no single, unified evaluative strategy being employed by readers to make judgments on ethical situations in a text, its fictional status, etc. Instead, there are several independent strategies at work that, dependent upon textual features like narrative perspective, as well as relevant background knowledge, may or may not begin to operate simultaneously with effect upon each other. Narrative perspective is one of the most basic and most powerful textual features in driving reader response and evaluations. What was seen in this study is that, for both fictional and non-fictional 1st person narratives, readers engage texts from a

default position which initially conflates narratorial and authorial empathic stance, but separates the actual identities of the narrator and the author. For fictional, 1st person texts, as the relevant background knowledge for the author's life increases, so too does the conflation of narratorial and authorial identity. This was not the case for the non-fictional 1st person text and it remains to be seen if that is a function of this interpretative community, or, rather, something that is dependent upon textual markers within non-fictional texts. It is interesting to point out that by the mere manipulation of narrative perspective, authors can manipulate the reader's judgment of actual authorial empathy. Readers of 3rd person texts automatically separate the author's empathic stance from that of the text, while readers of 1st person texts automatically conflate them. Additionally, while it is unsurprising to see that readers of 1st person texts increasingly conflate narratorial identity with authorial identity with an increase in relevant background knowledge, what is surprising is that readers of 3rd person narratives seem resistant to ever conflate narratorial and authorial identity. This finding has serious implications for how reader's cognitively identify and locate authorial identity and empathy within a text, particularly for 3rd person and non-fictional texts.

Furthermore, for basic aesthetic judgments, like that of skill in representation, as well as inferential judgments about author knowledge for subject matter via the text, what was seen is that reader's of 1st person texts primarily use textual markers to make those judgments, independent of any change in background knowledge for the author's actual life history, while readers of 3rd person texts making those judgments are far more likely to be influenced by knowledge of the author's life. As we have already seen for 3rd person texts, readers, by default, separate narratorial and authorial identity and empathic stance,

meaning that in order to judge how well the author knows what he or she is narrating and how effectively this is done, they often rely on background knowledge of the author's life as well as their own domain-related knowledge, among other things. What is troubling is the resistance to use background knowledge to help make these judgments, this time shown by readers of 1st person narratives. Readers of 1st person narratives may be aware that they are already conflating the author's identity and empathic stance with the narrator's, and, in order to make the most unbiased judgments of knowledge and skill, segregate all knowledge of the author from their own domain-related knowledge of the subject matter, and instead use that knowledge as the primary evaluative guide.

First, it is important to again note that these findings may, in fact, represent the strategies of a particular interpretative community. How specific the demographics of an interpretative community interact with interpretative and evaluative strategies is one of the unanswered questions that consilient literary studies needs to address. Is this population a total combination of its age, race, gender, sexuality, religion, geographic location, ideology, cultural background, status as untrained critical readers, social status (students), etc., or are there several factors among those which are of the most importance? For example, does training as a critical reader actually help predict what kinds of strategies are employed, more so than the attainment of any particular educational level? If there are differences that are dependent upon training as critical readers, are there also then differences across disciplinary lines, and if so, what does that suggest? Among the follow-up studies already planned, this is perhaps the most important of them all in order to develop a clearer understanding of the reading processes described here.

Another key element of this experiment is that the actual fictional status of the texts was an unknown in every condition, which is most often *not* the case in real life, where stories are usually approached either as fictional or non-fictional before reading. Does the knowledge of a text's status beforehand effect judgments of its status after reading? Is this something unique to the information era and the post-structuralist, post-modernist movements, where all texts are eventually treated as some sort of narrative? Is that claim itself true? If there is an effect of background knowledge on reader's interpretations, is there likely to be a level at which the amount of information begins to have less of an effect? For example, if proponents of literary interpretation claim that discursive accumulation adds to the understanding and appreciation of a text, is there a level at which that understanding seems to draw no benefit from additional interpretations? Could you test readers of a text with no literary interpretations and compare them with readers of several different numbers of interpretations to see what the effects are upon literary interpretation and appreciation? There are possibly even more basic assumptions that could be tested, for example, whether or not readers have an appreciable understanding of the difference between fiction and non-fiction, or story versus fact, two distinctions that would be central to providing evidence for the evolution of a "lie-detection" module.

As stated earlier in this chapter, this study is not only the first of several planned to investigate the relationship between background knowledge for an author's life and effects on reader response in various interpretative communities, it is also being presented as part of the data gathering for a larger, more evolutionarily-based hypothesis: that of the "lie-detector" hypothesis. In this case, working with an untrained population of readers

was ideal in that it would reflect the default, or evolved, capacities to detect "lies" or fictional status of texts, if such an evolved capacity exists, and is transferable to literary processing. The pre-experiment position was that if a "lie-detector" module does exist, it would have developed for detecting lies about evolutionarily salient cues. As fiction is typically not directly evolutionarily salient, it was believed that no such capacity for detecting the text's fictional status accurately would be found. Of the 143 respondents who participated in the study, 120 assigned a definitive fictional status to the texts they were asked to evaluate. Overall, respondents correctly rated their text's fictional status just below chance levels, at approximately 45%. However, of the 120 that assigned their text a definitive fictional status, 94 – approximately 78% of all respondents – decided that their text was non-fictional. This was seen in similar proportions for all texts, regardless of actual fictional status as well as narrative perspective. This can be taken as solid evidence that for a population of untrained readers there is no evolved capacity to correctly identify literary fictions, as was predicted. While a great deal of work remains to amass evidence in order to prove (or disprove) the existence of an evolved capacity for lie-detection, this study can be taken as solid evidence that even if that capacity does exist, its cognitive abilities do not readily transfer to the literary domain. In other words, readers have to be trained in order to become successful at discerning complex texts' fictional status.

Notes

1) There are any number of excellent sources which articulate various theories about
    the nature of the interaction between cultural and biological evolution. I would be
    remiss not to mention the following among them: Merlin Donald's *A Mind so
    Rare* and *Origins of the Modern Mind*, Boyd's and Richerson's *Culture and
    Evolutionary Process* and *Not by Genes Alone*, Cavilli-Sforza's and Feldman's
    Cultural Transmission and Evolution, Cycles of Contingency by Oyama, Griffiths,
    and Gray, Dan Sperber's *Explaining Culture*, as well as the work of researchers
    like E.O. Wilson, Robert Trivers, Gary Williams, John Tooby and Leda Cosmides,
    Steven Gangestad, Jeffry Simpson, David Buss, Robin Dunbar, Geoffrey Miller,
    Steven Mithen, Donald Symons, David Sloan Wilson, and Pascal Boyer.

2) Gould's argument is that while there are certainly structures and processes that are
    clearly adaptive, like hands and the ability to see, there are others that arise out of
    either a re-tooling from original functions, or complete coincidence from the
    development of those adaptations. However, there are a number of evolutionary
    researchers, Geoffrey Miller and Steven Pinker among them, who have argued
    that art and aesthetic processes are, in fact, evolutionarily *old* adaptations, usually
    by claiming that art gives an edge to its practicers not in selective fitness, but,
    rather, in reproductive fitness, much like the elaborate display put forward by a
    peacock's tailfeathers.

3) The entire field(s) of chaos theory, complexity, and emergence studies are another
    challenge to holists who believe that irreducibility equates with the impossibility

of scientific study. These fields exist precisely by demonstrating the phenomena which are not reducible nor explicable by the component elements, are still subject to mathematical and physical regularities.

4) There have been a number of literary theorists who have recently begun to hypothesize about the evolutionary origin of narrative, among them Joseph Carroll, Jonathan Gottschall and Brian Boyd. Stanislas Dehaene's *Reading in the Brain* provides an extremely well supported case for the development of reading as a general cognitive process as an exaptation of other visual discriminations which were evolved for.

5) The original study documenting the social-cheating-detection module originally appeared in *Cognition* 31, published in 1989, pages 187-276. While the evidence for a social-cheating-detection module is impressive, it is not conclusive, and it has been most vigorously challenged by Dan Sperber and Vittorio Girotto.

Supplemental Materials

All subjects were given the following instructions to read; the instructions were also read out loud prior to the materials being handed out to subjects:

Instructions

You have received one Introduction Sheet and one Text Packet, which you will read at your desks, taking up to as long as but no more than 30 minutes to read all the materials. Please read ALL materials carefully as anything and everything you receive may show up in the examination phase; do NOT write or make any markings of any kind on either the Introduction Sheet or Text Packet forms, nor are you allowed to take notes of any kind on any of your own materials. When you have finished and are comfortable you have

understood everything you have read, remain seated, raise your hand, and the experimenter will take all of the Text materials and give you the Examination Form. Answer all of the questions on the Examination Form carefully. When you are satisfied with your answers, return the Examination Form to the experimenter and then you may then leave, as you are finished.

Background Knowledge Forms were as follows

NBK: O'Brien

1. Author: Tim O'Brien; Title: "The Man I Killed"

IBK: O'Brien

2. Author: Tim O'Brien; Title: "The Man I Killed"; Tim O'Brien was born in Austin, Minnesota in 1946, a town with approximately 20,000 people, and which became a setting often used in his writing. When O'Brien was twelve, he and his family moved to Worthington, Minnesota, called "the turkey capital of the world," another locale that would later be central to his writing. O'Brien holds a Bachelor's in Political Science, attended but did not graduate from graduate school at Harvard, held an internship at the Washington Post, and currently holds the endowed chair at the Master of Fine Arts program at Texas State University – San Marcos. His writing has won both the National Book Award and the James Fenimore Cooper Prize for Best Historical Fiction.

RBK: O'Brien

3. Author: Tim O'Brien; Title: "The Man I Killed"; Tim O'Brien was against the Vietnam War, but reported for service and was sent to Vietnam with what

had become the infamous Americal division due to its involvement in the My Lai massacre in 1968, an event which resulted in the deaths of approximately 500 unarmed civilians, mostly women, children, and seniors. O'Brien's tour of duty was 1969-70. O'Brien's career as a reporter gave way to his fiction writing after publication of his autobiographical memoir about the Vietnam War, *If I Die in a Combat Zone.* His work contains actual details of the situations he experienced, and while that is not unusual, his conscious, explicit, and metafictional approach to the distinction between fiction and fact is extraordinary.

NBK: Sacks

4. Author: Oliver Sacks; Title: excerpt from *The Man who Mistook his Wife for a Hat*

IBK: Sacks

5. Author: Oliver Sacks: Title: excerpt from *The Man who Mistook his Wife for a Hat*; Oliver Sacks is the author of several best-selling books. His 1973 book, *Awakenings*, was adapted into an academy-award winning film of the same name in 1990 starring Robert De Niro and Robin Williams. He is currently a professor at Columbia University where he holds the title of Columbia Artist. He has been a member of the Academy of Arts and Letters since 1996, and in 1999 became Fellow of the New York Academy of Sciences, and in 2002 he became Fellow of the American Academy of Arts and Sciences. He is also a frequent contributor to the *New Yorker* and the *New York Review of Books*.

RBK: Sacks

6. Author: Oliver Sacks: Title: excerpt from *The Man who Mistook his Wife for a Hat*; Oliver Sacks is professor of neurology, psychiatry, and writing at Columbia University. He previously spent many years on the clinical faculty of Yeshiva University's Albert Eintsein College of Medicine. Sacks is the author of several best-selling books, including several collections of case studies of people with neurological disorders like amnesia, autism, musical hallucination, epilepsy, phantom limb syndrome, schizophrenia, and Alzheimer's disease. Most recently, Sacks and his book *Musicophilia: Tales of Music and the Brain* were the subject of an episode of the PBS series *Nova*. Sacks lives in New York, where he has practiced neurology since 1965.

NBK: McEwan

7. Author: Ian McEwan; Title: excerpt from *Saturday*

IBK: McEwan

8. Author: Ian McEwan; Title: excerpt from *Saturday*; Ian McEwan is a Booker Prize-winning English novelist and screenwriter. He has been shortlisted for the Man Booker Prize for Fiction numerous times, winning the award for *Amsterdam* in 1998. In 2008, *The Times* named McEwan as one of "The 50 greatest British writers since 1945." In 2008, McEwan publicly spoke out against Islamism for its views on women and homosexuality. McEwan, a self-described atheist, said that Christianity was "equally absurd," and later clarified his initially misrepresented views in a

personal blog that his comments were not directed towards Islam or Christianity per se, but "extremism" within both of these religions.

RBK: McEwan

9. Author: Ian McEwan; Title: excerpt from *Saturday*; Ian McEwan is a Booker Prize-winning English novelist and screenwriter. In 2008, *The Times* named McEwan as one of "The 50 greatest British writers since 1945." Prior to publishing *Saturday*, McEwan spent two years observing Neil Kitchen, Consultant Neurosurgeon and Associate Clinical Director at The National Hospital for Neurology and Neurosurgery in London. McEwan also thanked Professor Ray Dolan of the Wellcome Trust Centre for Neuroimaging, who he called "the most literary of scientists," for reading the transcript of *Saturday* and making appropriate neurological suggestions.

CONCLUSION

The lines have been drawn in the sand, the armies have gathered on the battlefield. Warning shots have been fired, each side testing the other's mettle. Soon the time will come for all-out war, and only one can emerge victorious... Okay, so I admit that at times in this project I've been a bit hyperbolic, often arguing if literary studies continues to operate within its current status quo, it's soon to be little more than an academic curiosity, a relic only rarely exhumed by scientists to demonstrate how *not* to conduct research. I've also at least implicitly suggested that I believe scientific study of literature and the arts can make progress in a variety of avenues where literary scholars have not. Hyperbolic or not, both of these claims are true. The stakes are indeed high in this "battle" for the future of literary studies, a battle that could help set a precedent for the study of art in general, determining how it is studied, where (within what academic disciplines), and by whom (scientists or humanists, or both in tandem). And make no mistake about it, scientists are already studying art in a variety of new and exciting ways.

Richard Gerrig's empirical work on literature in the branch of psychology called discourse processing is already mapping how real readers react to real texts (and parts of those texts), fictional and non-fictional. Neuroscientists like the eminent visual researcher V.S. Ramachandran as well as Semir Zeki have embarked on a course of study dubbed "neuroesthetics," which seeks to determine the material, neuronal laws that

govern human reaction to art (primarily visual art, but also music). Ramachandran has also proposed that human language, and our cognitive facility with metaphor, may be explained by mirror neurons and synesthesia, the phenomena where stimuli are experienced as having additional sensory qualities (numbers that are experienced always as the same colors, colors as the same tones, etc.). Within evolutionary psychology, one of the most heated ongoing debates is over whether or not art, in general, and fiction, in particular, is an adaptation, and, if so, what its adaptive value is, and, if not, why it is so pervasive in human culture[1]. All of these are premier scientific research programs backed with impressive technology and envious funding. The near future will be exciting for anyone interested in how the brain "does" art. Right now, many of these research programs *aren't* consulting with literary specialists, perhaps for the simple reason that most literary specialists would have little to contribute, primarily because of the methodological differences covered in Chapter One. As most literary scholars have little or no experience in developing scientific hypotheses and coupling those with experiments designed to test those hypotheses, they can rarely offer little more than anecdotal evidence. It would be a terrible shame if literary scholars were to miss out on being a part of the unraveling of the mystery of literature simply because of disciplinary training.

What has been presented here barely touches upon the true potential of consilient cognitive literary studies. Beyond neuroesthetics, empirical research on texts like Gerrig's coupled with the empirical work done in memory studies could help guide a pedagogical approach to teaching literature, one which would emphasize the most memorable aspects of texts through the most efficacious ways of presenting the material. A novel like James Joyce's *Ulysses* presents teachers with an almost overwhelming challenge of identifying

the relevant biographical, socio-historical, inter-textual, and stylistic information necessary in producing the desired level of comprehension and appreciation. Consilient cognitive literary studies could map the interactions between a reader and the author's biography, historical influences, inter-textual references, and style so that a teacher could plan the most effective way of presenting the novel based upon their own pedagogical goals. The study presented here is the first in a series ultimately designed not just to determine how background knowledge for an author's biographical data effects their reading of a text, but also how that knowledge might be best used in the classroom; do we teach our students about an author's life before they read her text, or after, or, perhaps, not at all? Empathic responses to literature and the closely related topic of ethical responses to literature have seen renewed interest within literary studies – Todd Davis' and Kenneth Womack's edited collection of essays, *Mapping the Ethical Turn*, and Chesire Calhoun's edited collection *Setting the Moral Compass*, are two excellent examples that include work by eminent literary scholars on empathic and ethical responses to literature. What is lacking in both of these collections is an engagement with the rapidly growing field of neuroethics (which is exactly what it sounds like: the study of the area's and activity of the brain in situations the utilize conscious ethical deliberation) as well as the immense wealth of study done on empathy in empirical, cognitive, and neuroscientific settings. Cognitive narratologists like David Herman are already working to identify the basic, discrete perceptual and conceptual elements of narrative, and suggest models for the complex and diverse brain processes that give rise to the robust experience of reading literature. Evolutionary models are being used to look at the "evolution" and "selection" process that gives rise to the canon of literary works, a survival of the aesthetically fittest,

as well as track the historical progress of literary response to narrative. Lisa Zunshine is leading the exploration of the relationship between Theory of Mind and our ability to comprehend literature, as well as how literary authors might intentionally manipulate that faculty to produce some of the aesthetic effects that we associate with high literature. As more becomes known about mirror neurons and their link to language, we can only imagine the rich insights which might develop between that area of study and literature.

Moreover, the near future will also see more art produced under the continually-evolving influence of contemporary neuroscience. As science progresses and changes, so too will the art that reflects it. Already there is an abundance of books, movies, and video games that could benefit from engagement with cognitive science. Less has been said about these other forms of narrative in this project simply because the scope of artistic phenomena is simply so vast, all of which, as products of a human brain embedded within human culture, are open to a consilient cognitive approach. It is my sincere hope that it will not only be literary critics and theorists who embrace this approach, but scholars of every form of art, and humanists in every discipline. There is simply nothing to be lost and much to be gained. For example, the *Iron Man* movies, James Cameron's *Avatar*, the *Transformers* franchise, *Inception*, and the *Matrix* trilogy, were all enormous blockbusters, and they each directly touch on issues of consciousness, embodiment, and perception. Film critics who engage these movies are just as bound to be consilient as are literary critics writing about cyberpunk fiction. And just as within literature, just because a movie doesn't thematically address ideas of consciousness or embodiment it doesn't mean that consilient cognitive criticism has nothing to say about the models of the mind and perception being represented in the film, nor the processes involved in producing and

viewing movies. In fact, short film clips are an experimental standard in many neuroscientific studies – it seems nothing short of ridiculous to ignore what scientists are finding out about what happens in our brains when we watch film.

Video games, the narratives they produce, and the interactive involvement with them, are almost an untapped area of critical research and interpretation. For an industry that has made more money than all of America's films each year since 2005, dwarfed the music industry by 2007, and globally surpassed the movie industry in 2008, and shows no signs of regressing despite the economic downturn, little mainstream academic attention has been devoted to it[2]. Questions about how the manipulation of perspective (point of view) in first-person games like the *Call of Duty* and *Halo* franchises versus the more traditional third-person perspective of games like *Fable* and the *Final Fantasy* franchise seem clearly resonant with traditional research areas within film and literary theory. *Heavy Rain* is a game based on film noir, with four characters each being controlled separately by the user, as they each try and piece together the identity of a serial killer. What makes the game unique, and well suited to consilient cognitive analysis, is its narrative structure. Almost very decision by the player has dramatic effects for the narrative: main characters can be killed, entire scenes can be missed, and the ending and its epilogue are entirely determined by the earlier actions taken within the game world. The overall fluidity of the entire plot is what makes this game, and the gaming industry, different from even other advanced narrative media like film. While the movie *Clue* was released with four alternate endings, everything else about the movie, in fact the vast majority, was the same. In *Heavy Rain*, it isn't only the endings that change, but the entire on-going story. While it is a well-known (but little studied) phenomenon that part of

literature's value is that a story is often quite received quite differently by the same person at different points in that person's life, this is a story that *literally* will be different each time it is played.

Team Ico is a game developer that has produced several critically praised and award-winning games, notably *Ico*, *Shadow of the Colossus*, and the forthcoming *The Last Guardian*. These games are well known for several design elements that lead the player to a high level of immersion in the game: non-interfering gameplay, focus on the story and setting, minimal dialog, bloom lighting (a technique that reproduces the effects of light being photographed by real world cameras), and key frame animation, which determines *what* movement the viewer will see at any point in the game. *Shadow of Colossus* plays as a puzzle game, leading the player-character through a series of battles with the Colossi, each having a specific weakness that must be exploited in order to defeat them. However, the end of the game, which necessarily entails the defeat (and death) of the last of the Colossi, turns back on the player-character, making the player wonder at the value of having murdered the last of a truly magnificent species. Video games like those produced by Team Ico could easily be seen as the frontier of aesthetic studies, combining aspects of literary and film theory with other aspects unique to the medium. If, and it is more likely a matter of when, the humanities decide to legitimate the study of video games, it goes without saying that that study would benefit from a consilient cognitive approach as much, or perhaps even more due to its very nature, as the study of literature and film. And, unsurprisingly, video games and virtual environments are another staple of experiments within the cognitive sciences[3].

Olaf Blanke's research into out-of-body-experiences used virtual reality; patients watched looped video of their own backs being stroked, projected in the virtual environment several feet in front of them, while simultaneously having their own backs stroked. The effect was that by feeling the stroke *and* seeing it happen synchronously somewhere else in space, they felt their own body position shift towards the area in space where the stroke was being projected by the virtual reality headset. EEG biofeedback "games" have been used for sometime already to help train cognitively different brains; for example, children with ADD can be trained to produce more beta (fast) waves in a game where colored bars represent more fast or slow brain waves. The end result: more beta waves means better attention. Since its inception in the early to mid 2000's, EEG biofeedback games have become more sophisticated, allowing simple planes to be "flown" by brainwave with a similar objective: good brain activity raises the plane over a wall in its path, while too little of the right activity results in a crash.

Indeed, as science and technology progress, more and more of what is happening in the real world will begin to sound like science fiction, and only those people trained to differentiate the two will be able to produce scholarship that isn't itself at risk of becoming fiction. In what is perhaps almost eerily reminiscent of the earlier discussion of posthumanism and cyberpunk fiction, monkeys have recently been trained to control robotic arms with their brains. Andrew Schwartz has trained immobilized rhesus monkeys to watch the movements of a robotic arm as it reaches for food, grasps objects in nearby space, and performs other basic motor commands within the monkeys' motor vocabulary. Using electrodes implanted within areas of the motor cortex that responded selectively to certain actions (recall the earlier discussion of mirror neurons and their

ability to code for a particular action and that action's intention), the monkeys were able to manipulate the robotic arms using their own motor cortices in order to successfully reach for and eat the food 61 percent of the time (Swaminathan 1). Other experiments by other researchers have used the monkeys' brain waves to "teach" a robot how to walk and to manipulate a remote robot arm several hundred miles away via camera feed. Underscoring all of these lessons: it was only possible to get the robotic limb to do what the monkey's embedded brain already knew how to do, not the reverse. This is not the "downloading" of consciousness that some posthuman literary critics have fantasized about, but the interfacing of the embedded and embodied brain with a technology that makes use of the same signals that our bodies do.

The future for consilient cognitive studies, whether of literature, painting, video games, or film, is "wider than the sky," as Gerald Edelman once described the limits of human consciousness. But a great deal of work remains to be done if consilient cognitive studies are going to find purchase within the humanities. However, it is well worth repeating that science is blazing ahead in producing reliable human knowledge, with or without the participation of those working within the humanities, and no topic is off limits to scientific study; if humanists think they own the rights to the study of all things aesthetic or ethical, or representative or empathic, they are sadly mistaken. Instead, by adapting our methodologies to become scientific and by keeping our scholarship consilient with contemporary knowledge being produced in the relevant scientific disciplines, humanists can become leaders in the production of progressive, scientific knowledge in their own areas of specialty. There is no longer a neat divide between the "two cultures," as C.P. Snow once described the gulf between scientific and humanistic

studies. Indeed, there have been several recent books coyly suggesting that there are now three cultures, the third culture being populated by scientists that study areas traditionally associated with the humanities[4]. Moreover, neuroscientific approaches are quickly opening avenues of scientific study into almost every traditional humanistic concern; there's already been an explosion of recent studies on neuroesthetics and neuroethics, and these fields are less than a decade old[5].

Literary studies stands to gain a great deal from a consilient cognitive approach. The study presented earlier here was selected for its ability to demonstrate how, for just one question, a consilient cognitive approach can unite literary theory produced from the New Critics to the New Historicists to the post-structuralists by testing hypotheses. Indeed, the study was designed as a specific test of a nexus of literary arguments, specifically the theses put forward by the New Critics like John Crowe Ransom, William Wimsatt and Monroe Beardsley, and Cleanth Brooks, and then the post-structural antitheses of Roland Barthes and Michel Foucault, and the evolution of that view into Stephen Greenblatt's New Historicist approach. Literary theory is replete with near-hypotheses that a consilient methodology can recognize, clarify, and test. For example, William Wimsatt's and Monroe Beardsley's classic New Critical essay, "The Intentional Fallacy," argued stridently that "the design or intention of the author is neither available nor desirable as a standard for judging the success of a work of literary art," and that this stance "entails many specific truths about inspiration, authenticity, biography, literary history and scholarship" (3). What specifically stands out is their attitude towards biographical data. Although they argue for judging a literary work "like a pudding or a machine," strictly based upon how well it works, they do seem to recognize that readers

seek to connect the thoughts expressed in a literary work with a person, whether the narrator or the author (4). Of this they say, "We ought to impute the thoughts and attitudes of the poem immediately to the dramatic *speaker,* and if to the author at all, only by an act of biographical inference" (5). What the process of biographical inference *is*, they fail to mention, but it is clear that even for the staunchest of New Critics, the author's biographical history is *thought to be* a powerful source of information that can directly effect the reading of a literary text. And there is the hypothesis that became the source of this study: knowledge of an author's biography can shape textual interpretation. The study here set out to begin to map the specific effects knowledge of an author can have on textual interpretation and did indeed show a significant effect for background knowledge of an author's biography for several kinds of textual evaluations. The results, if replicated across interpretative communities, will not only deepen our understanding of how we read and appreciate literature, but can also guide the pedagogy of teaching literature.

Moreover, as consilient literary criticism builds a foundation,it will inevitably begin to draw closer to other disciplines. While I have mentioned the work of Richard Gerrig, Marisa Bortolussi and Peter Dixon, Jemeljan Hakemulder, David Miall, and Jonathan Gottschall in relation to the study I have performed, I haven't even delved into the possible relation to other more general psychological phenomena; the two most probable suspects being domain related comprehension and schematic priming. Domain-related comprehension is a well known effect in psychology in which "[k]nowledge in a given domain . . . facilitates the acquisition of new domain-related information" (Chiesi 270). Is the introduction of biographical material effecting readers in a way that is similar to or can be explained better by findings about domain-related comprehension?

Schematic priming is a lesser known phenomena, but one that seems strikingly related to the effect found in the study presented here. An experiment by Owens, Bower, and Black used short multi-episode stories to test story recall. These stories were non-literary and borderline boring. Take the following excerpt as a typical example:

> Nancy arrived at the cocktail party. She looked around the room to see who was there. She went to talk to her professor. She felt she had to talk to him but was a little nervous about just what to say. A group of people started to play charades. Nancy went over and had some refreshments. The hors d'oeuvres were good but she wasn't interested in talking to the rest of the people at the party. After a while she decided she'd had enough and left the party. (186)

What Owens, Bower, and Black did was add an introduction to each story for half of the subjects. The introduction for this particular episode was the following: "Nancy woke up feeling sick again and she wondered if she really were pregnant. How would she tell the professor she had been seeing? And the money was another problem" (185). Unsurprisingly, and quite in line with what was found in the study presented here, was that the presence of the introductions had a sizable positive effect on readers' memories for the rest of the stories. What Owens, Bower, and Black suggested was that the introductions functioned as a organization schema for readers, allowing them to access a schema like "an unwanted pregnancy," and then read, comprehend, and remember the story based upon the expectations of that schema. It certainly seems possible that biographical data for an author's life acts as an organizing schema, a primer of sorts that guides the readers' expectations for the text while simultaneously enhancing their memory for the text.

The challenge of consilient literary studies is primarily methodological, but it is also one of focus. Literary scholars must learn to apply scientific methods to their subject, whether through the design of controlled experiments or the establishment of observational correlation through statistical analysis. The scope of this challenge is as wide as the scope of literary studies itself and includes textual manipulation, study of individual differences in single readers as well as interpretative communities, imaging studies on the activity in the brain while reading, and the mining of literary knowledge already produced for hypotheses that need empirical testing. However, this is also a change in the focus of literary studies. While the active interpretation of texts will always have a place within literary studies, it is not the central activity of the study of literature, and, in fact, opens up entirely separate questions that need studying: What are the cognitive processes involved in interpretation? What social factors influence interpretation? What individual factors? What are the aggregate patterns of interpretation at a given period? For a given text? Consilient literary studies would begin to focus on what literature is as a cognitive artifact, and how it is produced by embodied human minds embedded in their physical, social, and cultural environments, and how it is perceived, processed, remembered, comprehended, and interpreted by those same readers. It could ask questions about why certain literary forms, styles, tropes, images, and periods seem to endure and enchant, while others do not, and, more importantly, it could form hypotheses about these questions and test them. Consilient literary studies would necessarily bring literary scholars into close working relationships with scholars in other disciplines, promoting a true interdisciplinarity. There is a great potential gain in the

understanding and appreciation of literature through consilient literary studies, but right

now most of what consilient literary studies has to offer is just that: potential.

Notes

1. Evolutionary psychology has a much greater challenge in identifying cognitive

    mechanisms as adaptive than does evolutionary biology in identifying physical

    structures; while it is easy to see the adaptive value of a structure like the eye, it is

    much harder to see the exact adaptive value of art. That being said, Tooby and

    Cosmides have provided the best set of criteria to date used to identify cognitive

    adaptations. They include: 1) An adaptive target. This is a "description of what

    counts as a biologically successful outcome in a given situation" (73). If art, or

    fiction, is ever to be successfully identified as an adaptation, part of the argument

    must include an explanation of its biological (i.e. fitness or reproductive

    enhancing) relevance. 2) Background conditions. As human cognition evolved

    over the last several hundred thousand years, it is important to remember that an

    adaptation would have developed in response to pressures in "the ancestral world

    that [were] relevant to the adaptive problem" (73). Art and fiction cannot be

    explained as adaptations by appealing to their *current* cultural value, but, rather,

    must be explained in terms of the advantage they conferred to our evolutionary

    ancestors. 3) A design. This is perhaps the most difficult of all the criteria. While

    describing the design of the adaptation is straightforward, "a description of the

    articulated organization of recurrent features in the organism that together

    comprise the adaptation," it becomes difficult with respect to cognitive

    mechanisms in that the design "or even the existence . . . of a proposed

    information-processing mechanism is frequently unknown" (73-74). Moreover,

quite often "an appropriate functional description of a design in what one is trying to discover" (74). In other words, while biological adaptations often come in rather discrete forms such as an organ or limb, cognitive adaptations are rarely so easily defined. The task confronting an evolutionary psychologist arguing that fiction is an adaptation thus includes trying to describe the cognitive functions that shape and drive that adaptation, to say nothing of that adaptation's neurobiological implementation. 4) A performance examination. This is simply a description of "what happens when the proposed adaptation mechanistically interacts with the world" (74). 5) A performance evaluation. This is a description of "how well (or how poorly) the design, under circumstances paralleling ancestral conditions, managed to produce the adaptive target" (74). Criteria 4 and 5 require modeling, experimentation, and quantitative analysis to attempt to show that the adaptation actually produces the advantages being argued.

2. A search of the MLA database returned only 30 articles published in refereed journals, and only one in a flagship literary journal: Steven Jones' "Second Life, Video Games, and the Social Text," published in *PMLA*.

3. In fact, just since 2009, a search of PSYCINFO, the premier database of psychological journals, reveals 921 articles that involve "virtual reality," and another 363 for "video games," with experiments whose methods are conducted via virtual environments or through video game interfaces not included. That's just over one and half *professional* articles *within psychology alone* on video games and virtual reality published for every single day since 2009. A similar search through the MLA database reveals the striking disparity: only 28 articles,

books, and chapters on "virtual reality" have been published since 2009, and just 9 on "video games."

4. See John Brockman's *Third Culture: Beyond the Scientific Revolution*, David Pollock's and Ruth Van Reken's *Third Culture Kids: Growing Up Among Worlds*, and *The Three Cultures: Natural Sciences, Social Sciences, and the Humanities in the 21st Century* by Jerome Kagan.

5. These are growing fields and the time for getting in on the ground floor, so to speak, is now. Since 2001, there have been just under 250 articles published on neuroethics, of which just about *half* have been published in the last two years. The psychological study of aesthetics has long been flourishing field, but it too shows a recent increase in popularity; of the 3364 articles published on aesthetics within psychology over the last decade, 949 of them have been published in the last two years, approximately 28 percent. Compare these numbers to the wealth of studies published on memory, 66336 in the last decade, and 17500 in the last two years alone (26 percent), and you get the sense of not only how new these fields are, but how fast they are growing.

WORKS CITED

WORKS CITED

Abbott, H. Porter. *Diary Fiction*. Ithaca: Cornell U. P., 1984.

---. "Cognitive Literary Studies: The Second Generation." *Poetics Today*. 27:4 (Winter 2006) 711-722.

Attaway, Jennifer. "Cyborg Bodies and Digitized Desires: Posthumanity and Phillip. K. Dick." *Reconstruction: Studies in Contemporary Culture*. 4.3 (Summer 2004).

Baron-Cohen, Simon. *Mindblindness*. Cambridge: MIT Press, 1995.

Barthes, Roland. "The Death of the Author." Lodge, *Modern Criticism*. 145-172.

Bateson, Gregory. *Steps to an Ecology of Mind*. Chicago: U Chicago P, 1972.

---. *Mind and Nature*. Cresskill: Hampton, 2002.

Benjamin, Walter. "The Storyteller." Lodge, *Modern Criticism.* 10-29.

Bergonzi, Bernard. *Exploding English: Criticism, Theory, Culture*. Oxford: Oxford UP, 1990.

Blanke, Olaf, Landis, T., Spinelli., L., and Seeck, M. "Out-of-body experience and autoscopy of neurological origin." *Brain*. 127 (2004) 243-258.

Bould, Mark, and Sherryl Vint. "Of Neural Nets and Brains in Vats: Model Subjects in *Galatea 2.2* and *Plus*." *Biography*. 30.1 (Winter 2007) 84-104.

Bortolussi, Marisa, and Peter Dixon. *Psychonarratology*. Cambridge: Cambridge UP, 2003.

Bruner, Gerome. "The Narrative Construction of Reality." *Critical Inquiry* 18 (Autumn 1991) 1-22.

---. *Actual Minds, Possible Worlds*. Cambridge: Harvard UP, 1986.

Bukatman, Scott. *Terminal Identity*. Durham: Duke U P, 1993.

Campbell, Miranda. "Probing the Posthuman: Richard Powers' *Galatea 2.2* and the Mind-Body Problem." *Reconstruction*. 4.3 (Summer 2004) <http://reconstruction.eserver.org/043/campbell.htm> May 12, 2009.

Chiesi, H., G. Spilich, and H. Voss. "Acquisition of domain-related information in relation to high and low domain knowledge." *Journal of Verbal Learning and Verbal Behavior*. 18 (1979). 257-273.

Chodat, Robert. "Naturalism and Narrative, or, What Computers and Human Beings Can't Do." *New Literary History*. 37 (2007) 685-706.

Churchland, Patricia, and Terrence Sejnowski. *The Computational Brain*. Cambrigde: The MIT Press, 1992.

Clark, Andy. *Being There*. Cambridge, MIT P, 1997.

---. *Natural Born Cyborgs*. New York: Oxford UP, 2004.

---. *Supersizing the Mind*. New York: Oxford UP, 2008.

Clark, Andy, and David Chalmers. "The Extended Mind." *Analysis*. 58:1 (January 1998) 7-19.

Cohen, Patricia. "Next Big Thing in English: Knowing They Know That You Know." *The New York Times*. April 1, 2010. Section C, Page 1.

Cosmides, Leda. "The Logic of Social Exchange: Has Natural Selection Shaped How

    Humans Reason? Studies with the Wason Selection Task." *Cognition* 31 (1989).

    187-276.

Cosmides, Leda, and John Tooby. *Evolutionary Psychology: A Primer*.

    <www.psych.ucsb.edu/researchcep/primer.html> 8/25/2010.

Craig, A.D. (Bud). "How do you feel – now? The anterior insula and human

    awareness." *Nature Reviews Neuroscience*. 10.1 (Jan 2009) 59-70.

Crane, Mary Thomas. *Shakespeare's Brain*. Princeton: Princeton UP, 2001.

Crews, Frederick. "The Grand Academy of Theory." *Theory's Empire: An Anthology*

    of Dissent. Eds. D. Patai & W. Corral, New York: Columbia UP, 2005. 218-

    233.

Damasio, Antonio. *Descartes' Error*. New York: G.P. Putnam, 1994.

---. *The Feeling of What Happens*. New York: Harvest, 1999.

Danielewski, Mark. *House of Leaves*. New York: Pantheon, 2000.

Dawkins, Richard. "A Survival Machine" *The Third Culture*. Ed. John Brockman.

    New York: Touchstone, 1995.

Dehaene, Stanislas. *Reading in the Brain*. New York: Viking, 2009.

Delbanco, Andrew. "The Decline and Fall of Literature." *New York Review of Books*.

    46, November 4, 1999.

DeLillo, Don. *Mao II*. New York: Penguin, 1991.

Dennett, Daniel. *Consciousness Explained*. New York: Backbay, 1991.

Dick, Phillip K. "Do Androids Dream of Electric Sheep?" *Four Novels of the 1960's*.

    New York: Library of America, 1968. 431-608.

---. "The Three Stigmata of Palmer Eldritch." *Four Novels of the 1960's*. New York:

Library of America, 1964.

Dyens, Ollivier. "Cyberpunk, Technoculture, and the Post-Biological Self." *CLCWeb:*

Comparative Literature and Culture. 2.1 (2000).

<http://docs.lib.purdue.edu/clcweb/vol2/iss1/1> 9-3-2009.

Edelman, Gerald. *Bright Air, Brilliant Fire*. New York: Basic, 1992.

---. *Second Nature*. New Haven: Yale U P, 2006.

---. *The Remembered Present*. New York: Basic, 1990.

Edelman, Gerald, and Giulio Tononi. *A Universe of Consciousness*. New York: Basic,

2000.

Edmond, Jacob. "The Closures of the Open Text: Lyn Hejinian's 'Paradise Found'."

Contemporary Literature. 50.2 (Summer 2009) 240-272.

Ender, Evelyne. *Architexts of Memory*. Ann Harbor: U Michigan P, 2005.

Eugenides, Jeffrey. *Middlesex*. New York: Picador, 2002.

Fish, Stanley. "Interpreting the *Variorium*." Lodge, *Modern Criticism*. 287-306.

---. "The Crisis of the Humanities Officially Arrives." October 11, 2010.

<http://opinionator.blogs.nytimes.com/category/stanley-fish/> October, 29,

2010.

---. "Crisis of the Humanities II." October 18, 2010.

<http://opinionator.blogs.nytimes.com/category/stanley-fish/> October, 29,

2010.

Foster, Thomas. "The Reappearing Body in Postmodern Technoculture." *Contemporary*

Literature. 42.3 (Autumn 2001) 617-631.

Frith, Chris. *Making Up the Mind*. Malden: Blackwell, 2007.

Frow, John. "Never Draw to an Inside Straight: On Everyday Knowledge." *New Literary History*. 33 (2002) 623-637.

Fuller, David, and Patricia Waugh, eds. *The Arts and Sciences of Criticism*. New York: Oxford UP, 1999.

Gallagher, Shaun. *How the Body Shapes the Mind*. Oxford: Clarendon P, 2005.

Galvan, Jill. "Entering the Posthuman Collective in Phillip K. Dick's *Do Androids Dream of Electric Sheep?*" *Science Fiction Studies*. 24.3 (Nov 1997): 413-429.

Gerrig, Richard. *Experiencing Narrative Worlds*. New Haven: Yale UP, 1993.

Gibbs, Raymond W., Jr. *Embodiment and Cognitive Science*. Cambridge: Cambridge UP, 2006.

---. *Intentions in the Experience of Meaning*. Cambridge: Cambridge UP, 1999.

Gibson, William. *Neuromancer*. New York: Ace, 1984.

---. *Pattern Recognition*. New York: Berkley, 2003.

---. *Spook Country*. New York: Berkley, 2007.

Glimcher, Paul. *Decisions, Uncertainty, and the Brain*. Cambridge: MIT Press, 2003.

Goldman, Alvin. *Simulating Minds*. New York: Oxford U P, 2006.

Golumbia, David. "Resisting 'The World': Phillip K. Dick, Cultural Studies, and Metaphysical Realism." *Science Fiction Studies*. 23.1 (March 1996) 83-102.

Gottschall, Johnathan. *Literature, Science, and a New Humanities*. Houndsmills: Palgrave Macmillan, 2008.

Gould, Stephen Jay. *The Hedgehog, The Fox, and the Magister's Pox*. New York: Three Rivers, 2003.

Greenblatt, Stephen. "The Circulation of Social Energy." Lodge, *Modern Criticism*. 494-511.

Greenwall, Bill. "The Curious Incidence of Novel's About Asperger's Syndrome." *Children's Literature In Education*. 35.3 2004

Grossman, Lev. "If You Only Read One Mystery Novel This Summer." *Time Magazine*. 11 August 2003. <www.time.com/time/magazine/article/0,9171,10053963,00.html> 1 June 2009.

Guterson, David. *Our Lady of the Forest*. New York: Vintage, 2003.

Haddon, Mark. *The Curious Incident of the Dog in the Night-Time*. New York: Vintage Contemporaries, 2003.

---. Interview with Martha Woodruff. "Mark Haddon Discusses His Book *The Georgia Review Curious Incident of the Dog in the Night-Time*." *Fresh Air*. National Public Radio. 36 June 2003. <www.npr.prg/templates/story/story.php?storyId=4707767> 2 July 2009.

Hakemulder, Je'meljan. *The Moral Laboratory*. Philadelphia: Benjamin's, 2000.

Haney II, William S. *Cyberculture, Cyborgs, and Science Fiction*. New York: Rodopi, 2006.

Haraway, Donna. "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century." *Simians, Cyborgs, and Women*. New York: Routledge, 1991. 149-181.

Harrison, Kathryn. *The Seal Wife*. New York: Fourth Estate, 2002.

Hassan, Ihab. "Prometheus as Performer: Toward a Posthumanist Culture? A

University Masque in Five Scenes." *Georgia Review*. 31 (1977) 33.

Hayles, N. Katherine. *How We Became Posthuman*. Chicago: U Chicago P, 1999.

---. "The Posthuman Body: Inscription and Incorporation in *Galatea 2.2* and *Snow

Crash*." *Configurations*. 5.2 (1997) 241-266.

Herbert Jr., T. Walter. "Nathaniel Hawthorne, Una Hawthorne, and *The Scarlet Letter*:

Interactive Selfhoods and the Cultural Construction of Gender." *PMLA*. 103.3

(May 1988) 285-297.

Herman, David, ed. *Narrative Theory and the Cognitive Sciences*. Stanford: CSLI,

2003.

---. "Stories as a Tool for Thinking." Herman, *Narrative*. 163-192.

---. "Introduction." Herman, *Narrative*. 1-30.

---. "Spatial Reference in Narrative Domains." *TEXT*. 21 (2001): 515-541.

Hogan, Patrick Colm. *Cognitive Science, Literature, and the Arts*. New York:

Routledge, 2003.

Holland, Norman. "The Mind and the Book: A Long Look at Psychoanalytic Literary

Criticism." *Journal of Applied Psychoanalytic Studies*. 2 (2000) 13-23.

--- *The Critical I*. New York: Columbia UP, 1992.

---. *Literature and the Brain*. PsyArt Foundation, 2009.Documents

Hollinger, Veronica. "Cybernetic Deconstructions: Cyberpunk and Postmodernism."

Mosaic. 23.2   (Spring 1990) 29-44.

Hollinger, Veronica, and Joan Gordon, Eds. *Edging Into the Future*. Philadelphia, U
Penn P, 2002.

Iacoboni, Marco. *Mirroring People*. New York: Farrar, Strauss, and Giroux, 2008.

Jackson, Tony. "Issues and Problems in the Blending of Cognitive Science,
Evolutionary Psychology, and Literary Study." *Poetics Today*. 23:1 (Spring
2002) 161-179.

Jahn, Manfred. "'Awake! Open your eyes!' The Cognitive Logic of External and Internal
Stories." Herman. *Narrative.* 195-213.

Jameson, Fredric. *Postmodernism*. Durham: Duke U P, 1991.Documents

---. "Cognitive Mapping." *Marxism and the Interpretation of Culture*. Nelson, C. and
Grossberg, L, eds. U Illinois P, 1990. 347-360.

Johnson, Mark. *The Body in the Mind*. Chicago, U Chicago P, 1987.

---. *The Meaning of the Body*. Chicago: U Chicago P, 2007.Documents

Kakutani, Michiko. "Math and Physics? A Cinch. People? Incomprehensible." *New
York Times*. 13 June 2003.

Kaplan, Steven. "The Undying Uncertainty of the Narrator in Tim O'Brien's *The Things
They Carried*." *Critique*. 35 (1993): 43-52.

Kensinger, Elizabeth. "Human Memory: Once More, With Feeling." *Experimental
Psychology*. 54.3 (2007): 243-244.

Kermode, Frank. "Changing Epochs." Patai, *Theory's Empire*. New York: Columbia
UP, 2005.

Keysers, C., Wicker, B., Gazzola, B., Anton J.L., Fogassi, L., and Gallese, V. "A Touching Sight: SII/PV Activation During the Observation and Experience of Touch." *Neuron*. 42 (2004): 335-346.

Kirby, Vicky. *Telling Flesh*. New York: Routledge, 1997.

Klein, Kitty. "Narrative Construction, Cognitive Processing, and Health." Herman, *Narrative*. 56-84.

Koriat, Asher. "Metacognition and Consciousness." *The Cambridge Handbook of Consciousness*. Eds. Philip Zelazo, Morris Moscovitch, and Evan Thompson. Cambridge: Cambridge UP, 2007. 289-318.

Koriat, Asher, and Morris Goldsmith. "Monitoring and Control Processes in the Strategic Regulation of Memory Accuracy." *Psychological Review*. 103.3 (1996): 490-517.

Kuhn, Thomas. *The Structure of Scientific Revolutions, second edition.* Chicago: U Chicago P, 1970.

Lakoff, George, and Mark Johnson. *Metaphors We Live By*. Chicago: U Chicago P, 1980.

Laszlo, Janos. *The Science of Stories*. New York: Routledge, 2008.

Latham, Rob. "Mutant Youth: Posthuman Fantasies and High-Tech Consumption in 1990's Science Fiction." Hollinger, *Edging*. 124-141.

Ledoux, Joseph. *The Emotional Brain*. New York, Simon & Schuster, 1996.

---. *Synaptic Self*. New York: Viking, 2002.

Lim, Dennis. "The Life of Brain." *The Village Voice*. 15 March 2005. <www.villagevoice.com/books/0511bklim.62101.10.html> 24 July 2009.

Lloyd, Dan. *Simple Minds*. Cambridge: MIT Press, 1989.

Lodge, David, and Nigel Wood, eds. *Modern Criticism and Theory, 2ⁿᵈ ed.* New York:

Pearson, 2000.

Margolin, Uri. "Cognitive Science, the Thinking Mind, and Literary Narrative." Herman,

Narrative. 271-294.

Maturana, Humberto, and Francisco Varela. *The Tree of Knowledge*. Boston:

Shambhala, 1992.

McCaffery, Larry, Ed. *Storming the Reality Studio*. Durham: Duke U P, 1991.

---. "An Interview with William Gibson." McCaffery, *Storming*. 263-285.

McClure, John. *Partial Faiths*. Athens: U of Georgia P, 2007.

McEwan, Ian. *Saturday.* New York: Doubleday, 2005.

---. *Enduring Love*. New York: Anchor, 1997.

McHale, Brian. *Postmodernist Fiction*. New York: Routledge, 1987.

Menand, Louis. "Dangers Within and Without." *Profession*. 2005. 10-17.

Miall, David. *Literary Reading*. New York: Peter Lang, 2006.

Moretti, Franco. *Graphs, Maps, and Trees*. New York: Verso, 2005.

Nalbantian, Suzanne. *Memory in Literature*. Houndsmills: Palgrave Macmillan, 2003.

Obayashi, Shigeru, and Tetsuya Suhara, Koichi Kawabe, Takashi Okauchi, Jun   Maeda,

Yoshihide Akine, Hirotaka Onoe, Atsushi Iriki. "Functional Brain Mapping of Monkey

Tool Use." *Neuroimage*. 14.4 (October 2001): 853-861.

O'Brien, Tim. *The Things They Carried*. New York: Broadway, 1998.

---. "The Vietnam in Me." *New York Times Magazine*. 2 Oct 1994: 48-57.

---. *If I Die in a Combat Zone*. Seymour Lawrence: Delacorte, 1969.

---. *Going After Cacciato*. New York: Broadway, 1978.

O'Hara, Daniel. "Global Freud: Psychoanalytic Cultures and Classical Modernism."

     Journal of Modern Literature. 25.3-4 (Summer 2002) v-vii.

Owens, J., G. Bower, and J. Black. "The 'Soap Opera' Effect in Story Recall."   *Memory*

and Cognition. 7 (1979) 185-91.

Palmer, Alan. *Fictional Minds*. Lincoln: U Nebraska P, 2004.

Patai, Daphne and Will Corral, eds. *Theory's Empire: An Anthology of Dissent*. New

     York: Columbia UP, 2005.

Pence, Jeffrey. "The End of Technology: Memory in Richard Power's *Galatea 2.2*."

     Modern Language Quarterly. 63.3 (Sept. 2002) 343-363.

Pinker, Steven. *The Blank Slate*. New York: Viking, 2002.

---. *How the Mind Works*. New York: Norton, 1997.

Powers, Richard. *Galatea 2.2*. New York: Harper Perennial, 1995.

---. *The Echo Maker*. New York: Picador, 2006.

Rader, Margaret. "Context in Written Language: The Case of Imaginative Fiction."

     Spoken and Written Language. Ed. Deborah Tannen. Norwood: Ablex, 1982.

     185-198.

Ramachandran, V.S. *Phantoms in the Brain*. New York: Quill, 1999.

---. *The Tell-Tale Brain*. New York: W.W. Norton, 2011.

Richardson, Alan, and Ellen Spolsky, eds. *The Work of Fiction*. Aldershot: Ashgate,

     2004.

Richardson, Alan, and Francis Steen. "Literature and the Cognitive Revolution: An

    Introduction." *Poetics Today*. 23:1 (Spring 2002) 1-8.

Rizzolatti, Giacomo. *Mirrors in the Brain*. New York: Oxford UP, 2008.

Rorty, Richard. "Analytic philosophy and narrative philosophy." Lecture II, Pecs. 4

    May 2004.

Ryan, Marie-Laure. "Cognitive Maps and the Construction of Narrative Space." Herman,

    Narrative. 214-242.

Sacks, Oliver. *An Anthropologist on Mars*. New York: Alfred P. Knopf, 1995.

---. *The Man Who Mistook His Wife for a Hat*. New York: Harper Perennial, 1970.

Schacter, Daniel. *Searching for Memory*. New York: Basic, 1996.

Searle, John. *Mind: A Brief Introduction*. New York: Oxford U P, 2004.

Simon, Scott. "A 'Saturday' in the Life with Novelist McEwan." *Weekend Edition*

    National Public Radio. 19 March 2005.

        <www.npr.prg/templates/story/story.php?storyId=4538509> 2 August   2009.

Singer, Tania, Seymour, B., O'Doherty, J., Kaube, H., Dolan, R.J., Frith, C.D.,

    "Empathy for pain involves the affective but not sensory components of pain."

    Science. 303   (5661), 1157-1162.

Slingerland, Edward. *What Science Offers the Humanities*. Cambridge: Cambridge

    UP, 2008.

"Some Key Dates in Autism History." *Washington Post*. July 1, 2008.

    http://www.washingtonpost.com/wpdyn/content/article/2008/06/27/AR200806

    2703062.html November 4, 2010.

Stephenson, Neale. *Snow Crash*. New York: Bantam, 1992.

Stigler, Stephen. *The History of Statistics*. Cambridge: Harvard U P, 1986.

Swaminathan, Nikhil. "Monkey Think, Monkey Do – With a Robotic Arm."

    Scientific American. May 29, 2008.

        <http://www.scientificamerican.com/article.cfm?id=monkeys-see-monkey-do-

    wit>

Tabbi, Joseph. *Postmodern Sublime*. Ithaca: Cornell U P, 1995.

---. *Cognitive Fictions*. Minneapolis: U Minnesota P, 2002.

Toga, Arthur, and John Mazziotta, eds. *Brain Mapping: The Systems*. San Diego:

    Academic, 2000.

Tononi, Giulio, and Christof Koch. "The Neural Correlates of Consciousness: an

    update." *Annals of the New York Academy of Sciences*. 1124 (2008) 239-261.

Tooby, J. & Cosmides, L. "The Psychological Foundations of Culture**."** In J. Barkow,

    L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology*

    and the Generation of Culture. New York: Oxford University Press, 1992.

Turner, Mark. *The Literary Mind*. New York: Oxford UP, 1996.

Varela, Francisco, Evan Thompson, and Eleanor Rosch. <u>The Embodied Mind</u>.

    Cambridge: MIT P,   1993.

Vint, Sherryl. *Bodies of Tomorrow*. Buffalo: U Toronto P, 2007.

---. "Speciesism and Species Being in *Do Androids Dream of Electric Sheep*?."   *Mosaic*

40.1 (March   2007) 111-126.

Ward, Jamie. *The Student's Guide to Cognitive Neuroscience*. New York: Psychology

    Press, 2006.

Waugh, Patricia. *Metafiction*. New York: Routledge, 1984.

Weich, Dave. "The Curiously Irresistible Literary Debut of Mark Haddon."

   Powells.com 24 June 2003. <www.powells.com/authors/haddon.html> June

   17, 2009.

Weisbuch, Robert. "Six Proposals to Revive the Humanities." *Chronicle of Higher*

   Education. 26 March 1999, B4-5.

Whitehead, Colson. "Migratory Spirits." *The New York Times: Sunday Book Review.*

   October 22, 2006.

   http://www.nytimes.com/2006/10/22/books/review/Whitehead.t.html

   November 3, 2010.

Wilson, E.O. *Consilience*. New York: Vintage, 1998.

Wimsatt, Jr., W.K., and Monroe Beardsley. "The Intentional Fallacy." *The Verbal*

   Icon. Lexington: U of Kentucky P, 1954.

Winterson, Jeanette. *Sexing the Cherry*. New York: Grove, 1989.

Wood, James. "On a Darkling Plain." *New Republic*. 18 April 2005: 28-36.

Zunshine, Lisa. *Why We Read Fiction*. Columbus: Ohio State U P, 2006.

VITA

VITA

Louis J. Slimak, Ph.D.

EDUCATION ---

Purdue University, West Lafayette, IN, 2011
PhD in Literary Studies – English
Dissertation: *Consilient Cognitive Literary Theory* (Defended: September 14, 2011)
Supervised by: William Palmer, English, Co-chair; Ryan Schneider, English, Co-chair;
Paula Leverage, Foreign Languages and Literatures; Jeff Karpicke, Psychology
Primary Area: 20th Century American Literature
Secondary Areas: Cognitive Science, Cognitive Literary Theory, 20th Century British
Literature
Research Interests: Cognitive Studies, Science Writing, Science and Literature

The University of Akron, Akron, OH, 2007
MA in Literary Studies - English

The University of Akron, Akron, OH, 2000
BA in English, Honors, Summa cum laude

REFEREED PUBLICATIONS --

"The Transport of St. Ann, or How Many Angels Can Fit in Your Limbic System:
Neurology and Modern Faith in David Guterson's *Our Lady of the Forest*": *StoryTelling:
Journal of Popular Narrative*, 8:1, Summer 2008

CONFERENCE PRESENTATIONS --

"Consilient Literary Studies." The Louisville Conference on Literature and Culture Since
1900, 2011
"Consilient Literary Theory." Midwest Conference on Literature, Language, and Media,
Northern Illinois University, 2010
"Thinking All the Time: Preconscious Narratives in Ian McEwan's Fiction." Theory of
Mind Conference, Purdue University, 2007

"The Scientific Presentation of Consciousness in Ian McEwan's *Saturday.*" The Louisville Conference on Literature and Culture Since 1900, 2007
"The Melting Pot: Teaching Literature in the College Composition Classroom." MMLA Conference, Chicago, 2006
"Traumatic Space: Cognitive Mapping in Tim O'Brien's *The Things They Carried.*" University of Akron Graduate Symposium, 2006

AWARDS AND HONORS ---

Quintilian Award: Top Ten Percent of PICES Instructor Evaluations – Fall 2010 (ENGL 106)
Von's Bookshop Award for Literary Criticism 2009
Quintilian Award: Top Ten Percent of PICES Instructor Evaluations – Fall 2009 (ENGL 106)
Quintilian Award: Top Ten Percent of PICES Instructor Evaluations – Fall 2007 (ENGL 106)
Carl H. and Dorothy S. Bauer Scholarship 2000

TEACHING EXPERIENCE --

Purdue University, West Lafayette, IN
August 2007 – current: Full responsibility for curriculum creation, course design, text selection, lesson plans, daily class execution, grading, designating learning outcomes, conferencing

ENGL 250: Great American Books: American Consciousness (2 sections) – For this introductory level course, I emphasized basic narratological elements of literature and their functions within texts. Using fictional texts from Henry James, Thomas Pynchon, Phillip K. Dick, Richard Powers, and Don DeLillo, among others, as well as poetry by Billy Collins, e.e. cummings, and Emily Dickinson, philosophical essays by Ralph Waldo Emerson and V.W. Quine, and scientific essays by V.S. Ramachandran, the course considered the different ways Americans have represented consciousness over the last 150 plus years, and the unique place fiction has in that discourse.

ENGL 238: Introduction to Fiction: Cognitive Narratology (1 section)-
Another introductory course, the emphasis here was on identifying and understanding the basic narratological elements of fictional pieces, and how their manipulation effects readers. Texts from Toni Morrison, Margaret Atwood, Virginia Woolf, Don DeLillo, and Aldous Huxley, among others, were used to demonstrate a range of narratological stances, strategies, and styles. Guiding all of the course readings was the evolutionary question: What is the value of fiction in contemporary society and culture?

ENGL 108: Accelerated First-Year Composition (1 section) - In this honors version of the standard one semester composition course at Purdue, the traditional elements of argument

and Greek rhetoric provided the launching pad for various contemporary modes of analysis and composition, among them: digital rhetorics (including website design and blogging), visual rhetorics (film and photography), scientific analysis, political advocacy, and theological argument.

ENGL 106: First-Year Composition (9 sections) – This standard one semester composition course focuses on the traditional elements of argument and Greek rhetoric and places them side by side with contemporary arguments in a variety of media, modes, and contexts. From magazines to blogs, YouTube to CNN, textbooks to student essays, students are taught to evaluate each argument for its use of rhetorical awareness and savvy, and to produce their own compositions in a similar variety of media.

The University of Akron, Akron, OH
August 2005 – May 2007: Full responsibility for curriculum creation, course design, text selection, lesson plans, daily class execution, grading, designating learning outcomes, conferencing

ENGLISH 110: Introduction to English Composition (3 sections) – The first semester of two semester core requirement focused on teaching students basic elements of a successful written composition: theses and hypotheses, evidence and support, warrants, sentence structure and style, transitions, as well as developing their own personal stances on a variety of topical issues.

PROFESSIONAL AND DEPARTMENTAL SERVICE --------------------------------

Founding Steering Committee Member and Webmaster for The Center for Cognitive Literary Studies at Purdue University
August 2007 – current
- Design, development, and maintenance of university web page: <http://www.cla.purdue.edu/coglit/>
- Assistance in creation of certificate granting program within University
- Development of new classes
- Organization of public events: including travel arrangements and accommodations for guest speakers, site management, technical support, and advertising
- Management of bi-weekly public reading group

Syllabus Approach Leader for First-Year Composition at Purdue University
August 2008 – May 2009
- Evaluation and observation of incoming graduate instructors
- Review and approval of approach members' syllabi
- Development and implementation of department level policy, including approach criteria, rhetoric text approval, and assignments

Graduate Studies Advisor, Graduate Student English Association (GradSEA), 2009-2010

- ⮝ Worked with English Department Director of Graduate Studies and the Graduate Studies Committee to negotiate the inclusion of extra-departmental courses as qualified electives for several degrees within the department