

# Against Research: Literary Studies and the Trouble with Discourse

Paul Grimstad\*

*What we wanted was a machine that could learn.*

Alan Turing, “*Lecture on the Automatic Computing Engine*”

## 1. A Science of Literature?

Recently, in these pages, Eric Sundquist described what he called the “perilous” state of the humanities, pointing to the “avalanche of books about the crisis in liberal arts education in recent decades” (591). Indeed, “crisis,” as Sundquist acknowledges, has been the very condition of literary studies since its relatively recent beginnings. Already in H. C. G. Brandt’s 1884 keynote address at the first-ever meeting of the Modern Language Association (MLA), attendees were warned that if “teachers of modern languages . . . do not realize that their department is a science,” they are forced to conclude that “*anybody* can teach French or German or what is just as dangerous, any body can teach English” (58, 60; emphasis original). But Brandt assured his listeners that a “scientific basis dignifies our profession” (60), since English is the “historical scientific study of the language, Beowulf and Chaucer” (61) and that the practitioners of such a discipline “must be . . . specially and . . . scientifically trained” (60). While the context for Brandt’s claims about the scientific basis of English was the idea that modern languages could be approached with the philological rigor brought to the study of Greek and Latin, the first MLA address shows a basic anxiety about what it

\*Paul Grimstad is an assistant professor of English at Yale. His first book, *Experience and Experimental Writing: Literary Pragmatism from Emerson to the Jameses* came out in 2013. His writing has appeared most recently in the collections *American Impersonal: Essays with Sharon Cameron* (2014) and *Stanley Cavell and Literary Studies: Consequences of Skepticism* (2011).

*American Literary History*, vol. 26, no. 4, pp. 647–663

doi:10.1093/alh/aju050

Advance Access publication October 8, 2014

© The Author 2014. Published by Oxford University Press. All rights reserved.

For permissions, please e-mail: journals.permissions@oup.com

means to think of literary studies as a science. Louis Menand has described this as arising from the “incorpora[tion of] literature into the structure of the research university,” and so of giving the study of literature a “sciencelike status” (109). While the sense of crisis would persist over the next century in a series of atavistic resurfacings—researcher versus generalist; scholar versus critic; formalist versus historicist<sup>1</sup>—what is perhaps genuinely new in the latest wave of what Menand calls literary studies’s “obsessive” “self-examination” (61) is the shift from questions about *how* literature should be studied to more pointedly skeptical questions: just what *is* it that disciplines like English or comparative literature or philosophy teach? What are the statuses of claims in these disciplines and what are the criteria for how evidence is used to evaluate them? What problems are they inheriting and trying to solve? What are their objects of analysis? What exactly constitutes research in these disciplines?

One of the more provocative instances of literary studies’s self-examination—Steven Knapp and Walter Benn Michaels’s 1982 essay “Against Theory”—claimed that if “theory” names in part the attempt to arrive at general principles for the practice of literary criticism, then to be “against” it is to be against the idea that there could be a metalanguage that would prescribe methods for criticism from a standpoint outside practice. Knapp and Michaels treat the question of whether literary criticism needs a theory of practice as closely linked to another question central to scholarship across the humanities: understanding what goes on when one intends something or looks for intention in another’s sayings or actions. If the relation of theory to practice is bound up with the question of intention, then this invites a further set of questions about language and meaning: are meanings timeless linguistic types or is meaning always context-sensitive? Are persons required for there to be meaning or can meanings impersonally circulate? Could there be such a thing as an “intentionless meaning” (727)?

Thirty years later, in an essay declaring himself “sympathetic to the aims of and the argument of ‘Against Theory,’” Charles Palermo notes “how unavoidable the debate about intentionalism is,” saying that intentionality is “not an issue one may take a pass on.” Palermo raises the issue of intention in relation to what he calls the “anti-intentionalist default position” in literary studies, tracing this back to two essays which had an immense influence on the study of literature in the US and which are touchstone writings for much of what would come to be called “theory”—Roland Barthes’s “The Death of the Author” (1968) and Michel Foucault’s “What Is an Author?” (1971). Barthes’s essay questioned the idea that, in reading a work of literature, one looked for or even assumed the presence of authorial intention, since the idea of an author both presupposes a godlike

originating authority and is a historically contingent formation (1467). Foucault's essay generalized the move from the author to what he called the "author function." The author function, for Foucault, serves to "characterize the existence, circulation and operation of certain discourses within society" (124). Foucault had previously offered a more systematic account of the relation of author function to discourse in *The Archaeology of Knowledge* (1970) where he both clarified what he called the "rules of formation" of discourse and generalized the author function to account for artworks described as "nodes in a network" of a discursive formation (38, 23). Barthes himself went on to generalize all of this further in his essay "From Work to Text" (1971) identifying what he called a "mutation in the humanistic disciplines" involving nothing less than a "new object" 'unobjet nouveau' within the "field of the sciences" 'dans le champ des sciences' ("From Work" 1478; "De travail" 908). Here the move is from the "work" understood as an intentionally structured unity to "the Text" which, like the impersonal circulation of discourse, reduces discretely authored works into a "methodological field" (1478). The immense influence of these essays on literary studies in the US suggests that terms like *discourse*, *author function* and *the Text* satisfied a need for specialized objects required for the constitution of a field of research.

To claim these essays furnished literary studies with specialized objects of analysis suitable for a research program is to go against the idea, so pervasive during the culture wars, that the upshot of such thinking amounted to some form of relativism.<sup>2</sup> On the contrary, the most conspicuous features of theory and its specialized vocabulary circa 1970–1971 were its effort to give literary studies a science-like status. As Jeffrey Williams notes, such essays led the humanities to "measure themselves on what I call 'the research protocol,'" as disciplines "expected to produce new knowledge through research [with] metrics of evaluation [that] emulate those of the sciences" ("Rise" 693). If theory served the aims of coordinating disciplines in a "reach toward the scientific," this then made literature "more apt for research," subsequently giving rise to the theory journal (695). And since the theory journal was, as Williams notes, "a showroom of research," criticism came to imagine itself as the science of "uncovering the structural operation of language, interpretation, gender and society" ("Little" 408). Indeed, Foucault and Barthes's linked ideas of the methodological field of the Text, the author function, and the treatment of artworks as nodes in a network of discourse circulation were remarkably effective in providing general terms for a discipline looking to legitimate itself by giving precise definition to its objects of analysis.<sup>3</sup>

One could of course see Knapp and Michaels's argument against theory as a critique of the scientism inherent in the idea that literary criticism is a form of research, since a science of criticism would presumably require a theory of practice in relation to which critics may make "claims about the nature of [the] object" of literary study (736).<sup>4</sup> Indeed, for Foucault, such theory takes the form of thinking of authorship in functionalist, and artworks in operationalist, terms; for Barthes, it takes the form of positing a "new object" within the "field of the sciences." But what exactly would it mean to think that authorship and artworks—phenomena bound up unavoidably with human personhood—were explicable according to these sorts of metrics? The question becomes vivid, I want to argue here, when considered in relation to the question of whether personhood may be embodied in machines or modeled in a set of programmed instructions. In what follows, I elaborate an intellectual history beginning with Charles Peirce's work on logical machines and inference, moving to Alan Turing's laying the foundations for a philosophy of artificial intelligence (AI), then to Hubert Dreyfus's critique of the AI research program, and concluding with Knapp and Michaels's asking, in the context of their argument against theory, that we consider the question, "Can computers speak?" (729). In each case, the question at stake is whether personhood is reducible to the operation of a set of rules and thus subject to theoretical formalization; or, conversely, "whether computers are capable of intentions . . . [and] whether computers can be intentional agents" (729). With these questions in mind, I go on to show how a thinking of literary studies as a discipline "expected to produce new knowledge through research with metrics of evaluation that emulate those of the sciences" is bound up with a nest of confusions around the word *discourse*. If the dominant use of *discourse* in literary studies remains Foucault's influential account, his functionalism about the discursive has, I argue, a number of telling affinities with the view that human intelligence is reducible to the operation of a set of rules. Against this now completely orthodox understanding of discourse, I argue here that the discursive is more accurately—and, for the humanities, much more fruitfully—understood as that which makes possible the way persons use concepts in judgment and intentional action and take things to be thus-and-so: just those interpretive mental acts that are the end and object of scholarship in the humanities. I thus pursue in this essay an idea of human persons as *constitutively* discursive—understood in the sense of what computers can't do—rather than explicable according to a general theoretical object called "discourse."

## 2. Abduction, O-Machines, and What Computers Still Can't Do

While Brandt was assuring his fellow modern language scholars that their discipline can and must be practiced as a science, Peirce was teaching courses in logic at the first modern research university in the US. Building on the work of Allan Marquand, his student at Johns Hopkins (where he also taught John Dewey and Thorstein Veblen) Peirce wonders in an 1887 essay whether a mechanism might be designed that could carry out deductive inferences and, if so, whether such machines might be a steps toward other machines that would be capable not only of calculation but also of choice, hypothesis, creativity, and learning. If in a deductive syllogism, the relation among premise, case, and conclusion is one of necessity, then, Peirce believed, in principle a mechanism could be assembled that embodied that relation. But a reasoning machine, Peirce goes on, is “destitute of all originality. . . . [I]t cannot find its own problems, it cannot feed itself. . . . [T]he machine is utterly devoid of original initiative and would only do the special kinds of things it had been calculated to do” (“Logical Machines” 168–69). In a definition he later contributed to the *Century Dictionary*, Peirce wrote that the “value of logical machines seems to lie in their showing how far reasoning is a mechanical process, and how far it calls for acts of observation” (“Logical Machine,” def. 8). In beginning to think of reasoning as a nonmechanical process bound up with observation, Peirce imagined it as a particular sort of attention, one in which there is always a “nondeterministic step, which often involves creativity [and] discernment” (Ketner 49). He is wondering if there could be a machine that would be capable of absorbing new experience into its inferences—a machine that could learn.<sup>5</sup>

Later Peirce codified this expanded idea of reasoning in the type of inference he called “abductive.” Not limited to deductions that could in principle be carried out by a machine, abductive inferences are starting points which “provide hypotheses for testing” (Misak 18). Abduction is a conjecture at what some curious observation or circumstance “*might* mean. . . and testing inductively to see if these consequences are borne out in experience” (Anderson 47). Peirce says such abductive starting points come to us like a flash and arise from a choice or preference for one hypothesis over another; the postulating of a hypothetical rule that may or may not go on to explain some perceived particular (“Pragmatism” 226). Abduction, then, could be described as a conjecture at a general category that may account for some particular or set of particulars, not through the determinate calculations of deduction, nor through observed regularities that allow for inductive prediction, but as the choosing of a

starting point from which to begin reasoning. The ability to select such starting points requires capacities Peirce believed to be impossible to build into a machine.

Insights about the relation of logic, machines, and human intelligence similar to Peirce's were arrived at through different channels in Alan Turing's 1936 paper "On Computable Numbers with an Application to the *Entscheidungsproblem*." Here, Turing took up aspects of German mathematician David Hilbert's program to secure the foundations of mathematics. Like Bertrand Russell and A. N. Whitehead in their *Principia mathematica* (1913), Hilbert wanted to derive the whole system of arithmetic from logical principles. One of the questions arising from this work was what came to be known as the *Entscheidungsproblem*, or "decision problem." Hilbert wanted to find out if "there existed a definite method which could, in principle, be applied to any [mathematical] assertion, which was guaranteed to produce a correct decision as to whether the assertion was true" (Hodges 91). Essentially, what Hilbert wanted (though he did not put it this way) was a machine that could determine whether a given number was computable. Either the machine will come to halt or it won't; either the number is computable—decidable—or it isn't. To put it in these terms is to cast it in the image of a Turing machine, one of the central ideas in "On Computable Numbers." A Turing machine is not a piece of functioning hardware but a thought experiment involving an infinite tape divided up into discrete sections and a read-write head printing or removing symbols on the tape based on the state the machine is in at a given moment. Turing's idea was that to arrive at an answer that satisfied Hilbert's desire for decidability would require the kind of operation that could in principle be carried out by such a machine.

But Turing's work on the decision problem led him to a more speculative question: could we or should we think of human thought as comparable to what a Turing machine does? And he became convinced, as Peirce had earlier, that what a human person does when he "thinks" is not something that could be embodied by a machine carrying out programmed instructions. This conviction led Turing to imagine a new kind of machine that would "proceed deterministically, step-by-step, but once in a while make nondeterministic leaps, by consulting 'a kind of oracle as it were'" (Dyson 252). Such a machine would have a random step or guess built into its operations and so would be like a Turing machine, except that the machine "is endowed with an additional operation of a type that no Turing machine can simulate," one that "works by 'some unspecified means'" (Copeland 142). Turing believed his Oracle machines, or "O-Machines," were "closer to the way intelligence (real and artificial) works" (Dyson 252), since they would be, as he put it, "machine[s] that c[ould] learn from experience" (qtd. in Dyson 261).<sup>6</sup>

What Peirce called “abduction” and what Turing named an “oracle”—nondetermined jumps that are themselves a part of thinking—name the way each came to believe that human intelligence involves a continuously responsive agency that is very hard to get a machine to emulate. From the vantage of Turing’s innovations in mathematical logic and machine intelligence, we can see that Peirce had already thought through the elements for a nondeterministic thinking machine along the lines of an O-machine. But Peirce failed to put the pieces together. Those pieces were his idea of a machine that would carry out deductive inferences and the idea of abduction as the logic of discovery. What is most interesting and relevant here, though, is the way Turing’s thinking about the *Entscheidungsproblem*—and his later more explicit concern with AI in his 1950 paper “Computing Machinery and Intelligence”—bears out Peirce’s idea that thinking is something other than calculation.

That a machine following rules could learn to abduct starting points from which to begin reasoning is just the idea Dreyfus rejects in his book *What Computers Still Can’t Do* (1972). Whether the topic is game-playing, pattern recognition, or competence in a natural language, what the early AI advocates had failed to show, according to Dreyfus, is how these sorts of abilities can be reduced to a set of programmed instructions (72). As Dreyfus puts it, the research program that took off from “On Computable Numbers” and “Computing Machinery and Intelligence” sought to arrive at a technique “for finding the rules which thinkers from Plato to Turing assumed must exist—a technique for converting any practical activity . . . into the set of instructions . . . called a theory” (74).<sup>7</sup> But, Dreyfus notes, “even in logic,” such rules “are not around to be found because they do not exist independently of the pragmatic context” (119). This claim, which guides all of his criticisms of the AI research program, Dreyfus identifies with “what . . . Peirce called abduction” (21). If, in abducting, we are finding starting places from which to *begin* reasoning, then it is just this capacity a computer doesn’t have, since the AI model draws its inferences from a starting point decided in advance (for example, from programmed instructions). As Dreyfus puts it: “a timeless, contextless theory of competence *cannot* be used to reproduce the moment-to-moment involved behavior required for human performance. . . . [I]ndeed there cannot be a *theory* of human performance” (190–91); “[s]ince a human being uses and understands sentences in familiar situations, the only way to make a computer that can understand actual utterances and translate natural language may well be, as Turing suspected, to program it to learn about the world” (109).

What would it mean for machine to “learn about the world”? To make the question concrete, Dreyfus turns to a discussion of the

particular sort of competence that is knowing a natural language. For Dreyfus, when one learns a language, “[s]ome form of thinking other than searching, or counting, “is taking place,” a “function of intelligence” that is an intuitive ability to ferret out “the essential from the inessential,” such that only a machine capable of learning—that is, of being initiated through training into the practices that make up a form of life—could become competent in the use of a natural language (114, 119). Accordingly, competence of this sort is not a formalizable calculus, since there “there can[not] be a formal theory of what linguists call pragmatics” (198). Here, Dreyfus enlists Ludwig Wittgenstein’s later thought to make the point that in order to learn a natural language, a computer “must not only have grammatical and semantic rules but further rules which could enable [it] to recognize the context in which the rules must be applied” (203).<sup>8</sup> This regress of rules leads Dreyfus to conclude the impossibility of a formal theory of competence, since such a theory would “have to [be] a theory of all human knowledge” (198). In order to ward off the regress, there would need to be something like abduction: the sort of intentional agency that would allow both for the selection of a starting point from which to begin reasoning and for the skillful and assimilative undergoing that is learning.<sup>9</sup>

Dreyfus’s brief against AI hooks up in obvious ways with the argument of “Against Theory.” For Knapp and Michaels, meaning, interpretation, and intention presuppose one another since, as they put it, “language has intention already built into it” (736). The idea can be expressed using Dreyfus’s language: to interpret is to have picked out something essential from a background of inessentials. One does not interpret the meaningful utterance or mark in the discovery on a beach of what appear to be stanzas from a poem but are really the effects of erosion, according to the example from “Against Theory” (727). If the marks on the beach have not been left there intentionally—if they are just natural accidents—then they are not lines of a poem, or language at all, but merely resemble language. To interpret something *as* lines from a poem is to engage in an activity in which you’re never asking the question about the presence or absence of intention, since once you’ve decided that something warrants the sort of attention called “interpretive,” the question of intention is moot.

The affinity between Dreyfus’s critique of AI and “Against Theory” becomes most explicit when Knapp and Michaels take up the example about “how difficult it is to imagine a case of intentionless meanings” as captured in the question, “Can computers speak?” (727, 729). Because, they say, there is no such thing as an intentionless meaning, “the only real issue is whether computers are capable of intentions” (729). Could a computer, on any account, ever be thought of as intending something? To think so would commit one to



the idea of formal correctness (rule-following as the carrying out of a calculus) as the mark of intention. But as we have seen in Dreyfus's appeal to the later Wittgenstein, such an idea of formal correctness leads to a regress. This is because there is more to intention than operating a set of rules, namely, our holistic sense of what is essential and inessential. Knapp and Michaels's arguments about intention and Dreyfus's thoroughgoing critique of the AI research program are thus analogous in the following way: the reason we should not expect a computer to be able to intend things is that intentional agents get around in the world without a regress of rules; that the involved, moment-by-moment coping that is human personhood is not formalizable in a theory.

*[T]he reason we should not expect a computer to be able to intend things is that intentional agents get around in the world without a regress of rules.*

### 3. The Trouble with Discourse

While Dreyfus has himself been one of the most astute interpreters of Foucault in the US, he has distanced himself from just those aspects of Foucault's thought that have contributed to what Palermo called an "anti-intentionalist default" in disciplines like literary studies and art history. In the book he coauthored with Paul Rabinow (*Michel Foucault: Beyond Structuralism and Hermeneutics* [1983]), Dreyfus de-emphasized the Foucault of discursive formation, subject position, author function, and so on, because this was, as David Hoy puts it, "an aberration" in Foucault's thinking "[c]oinciding with the wave of structuralism in France in the Sixties. . . [which] stressed the analysis of discourse." One explanation of why Dreyfus might be tempted to see this side of Foucault as an aberration is that his critique of the AI research program works equally well as critique of the idea that personhood can be explained as subject position in a discourse formation. If, as Dreyfus argues, human personhood is not reducible to the operation of a set of rules—it is rather something like the abductive ability to ferret out essentials from a background of myriad inessentials—then the trouble with *discourse* is that it has become a reified term of art for work in the humanities that explains personhood as a nodal point in a network. This aspect of Foucault's thought comes through forcefully in his 1970 Collège de France lecture "The Discourse on Language," where he describes the "fact that people speak" as a "barely imaginable power" ("des pouvoirs. . . qu'on imagine mal" [*L'ordre* 10]) arising from "material manifestations" which "do not belong to us" (*Archaeology* 216).

In this last section, I want to argue, on the contrary, that nothing belongs to us more intimately than discourse, since it is the condition for our ability to abduct starting points from which to begin

reasoning, ferret out the essential from the inessential, take things to be thus and so, and intend things. What exactly does it mean to say that, rather than a position in a discourse network, persons are constitutively discursive? To stick with Dreyfus's example of competence in a natural language and the accompanying appeal to Wittgenstein on rules, consider an oft-cited passage from Stanley Cavell's 1962 essay "The Availability of Wittgenstein's Later Philosophy." Drawing on ideas central to the *Philosophical Investigations* (1951), Cavell describes our ability to "project words into new contexts" as a matter of our "sharing routes of interest and feeling, modes of response, senses of humor and of significance and fulfillment, of what is outrageous, . . . of when an utterance is an assertion, when an appeal, when an explanation—all the whirl of organism that . . . [is] 'forms of life'" (52). This account of what it means to be able to go on in a language game makes vivid how *discourse* in Foucault's sense misses something crucial about the direction of fit between persons and world. Rather than neutral receptacles confronted by a vast edifice of material circulation that explains (even determines) them, persons just are the array of capacities that buoy and sustain the "routes of interest"—the deep agreements and attunements—that constitute a form of life. The discursive, we might say, is the whole messy motley of all the things we do with words and the corresponding ability to use words intelligibly in unforeseeable situations. Were this to be imagined as a formal theory (as, for example, in functionalist language that would explain personhood as "subject position in a discourse network"), it would need to be a theory of all human knowledge, since it would reduce the motley of our practices—"all the whirl of organism"—to a general object.

A look at another recent exchange about the vicissitudes of the humanities's self-conception gets into still clearer view the difference between Foucault's account of discourse (at least around the time of the methodological treatise *Archaeology of Knowledge*, the essay "What is an Author?" and the 1970 Collège lecture, "Discourse on Language") and the account I have been pursuing. Like Dreyfus, Geoffrey Harpham recognizes the immense influence and importance of Foucault's work for fields as diverse as literary studies, legal theory, history, political theory, and anthropology, but nevertheless thinks Foucault's account of "anything that happens in the human realm" to be "singularly impoverished" (28). This makes Foucault's far-reaching influence on the humanities a problem, since, as Harpham puts it, "the humanities are inconceivable without some idea of the human," a claim he links to the idea that scholars in the humanities ought to "contemplate not only a proposition but the proposer" (26). Jonathan Culler objects to this account of the humanities and to the significance attributed to Foucault's influence on its

current state claiming that the appeal to “what human beings believe they are doing or intending to do is not a reliable guide to what is actually happening in history [and] in discourse” (39). Further, Culler notes that in persisting to refer to the body of scholarship done in disciplines like literary studies and philosophy as “the humanities,” we risk succumbing to an “ideology of humanism” (39).

Harpham and Culler’s disagreement over what the humanities are for (even over the appropriateness of the name) is relevant to arguments I am making here. The idea that humanistic inquiry ought to involve trying to construe the proposer behind the proposition is the basic claim of “Against Theory”: to look for meaning just is to look for intention since without a proposer there wouldn’t be anything recognizable *as* a proposition. But Culler seems to be saying that to be interested in personhood as intentional is to succumb to an ideological mystification (humanism); and, further, that analysis of *discourse* (presumably in the sense of discourse networks or discursive formations) will serve as an antidote to such mystification. But when Culler writes that the appeal to what human beings “believe they are doing or intending to do is not a reliable guide to what is actually happening . . . in discourse,” he has in mind an extremely narrow idea of intention. If by intention is meant the fact that persons are able to take things *as* things—the *überhaupt* aspect of intentionality at play in any autonomous language game—this has nothing whatsoever to do with what some individual person might “believe they are doing or intending,” as if intentionality must entail perfect transparency in our assessments of our own or another’s sayings or actions. That persons are discursive through and through is rather a claim about the holistic intelligibility of a form of life, and of a person’s capacity for taking it that meanings are in play regardless of what she may end up believing. Furthermore, would not susceptibility to ideological mystification (of any stripe) itself *depend* on capacities that make persons responsive to the world in just this holistic sense of intentionality? How otherwise would one recognize a set of beliefs to become mystified by at all? In any case, it would be very strange to claim that intentionality were intrinsically mystificatory. And lastly, and in a more meta-disciplinary register, if it is a mere “piety” to persist in caring about the intentionality of persons, what is to prevent one from treating Culler’s words as themselves issuing from a discipline anxious to justify what counts for it as an object of analysis (41)? Might we hear in the warning about humanist ideology (lurking in the word *humanities*) a predictable response from a discipline eager to define itself as research upon specialized objects like *Text*, *discursive formations*, *author functions*? Indeed, this is just the sort of caution we might expect from a literary studies concerned, as Williams put it, “to produce new knowledge through research [with]

metrics of evaluation [that] emulate those of the sciences,” with *discourse* thought of as providing those metrics.

If part of what theory gave to literary studies were generalized objects like discourse, and codified methodological protocols like “discourse analysis” (in the way Foucault taught whole generations of US academics to write and think), these lessons have contributed, I have been saying, to an idea of literary studies as a research program. That this is a category mistake becomes still clearer when we consider some of Thomas Kuhn’s insights about theory change in the natural sciences. Kuhn’s well-known idea of “paradigm shift” was his way of describing how a scientific community’s conception of its legitimate problems may change when perceived anomalies within the practice of “normal, puzzle-solving” science themselves come to be seen as law-like. As Kuhn puts it, these are “significant shifts in the criteria determining the legitimacy both of problems and of perceived solutions” (*Structure* 67, 108–09). In a later essay, “Objectivity, Value Judgment and Theory Choice,” Kuhn replied to critics who saw his discussion of paradigm shift as reducing the activity of scientists to “mob psychology” (321). Describing how a scientific theory arises from out of historically specific and overlapping sets of evaluative criteria Kuhn considers instances when there appear to be “at least some good reasons for [a number of conflicting] possible choices” between theories (360–61). A hoped for solution to such impasses is that “further research [will] eliminate residual imperfections and produce an algorithm able to dictate rational, unanimous choice” (359). Because Kuhn does not believe there could be such an algorithm, he concludes that the resolution of conflicting criteria for theory choice comes down to the “decisions of individual scientists” and the “scientists who share the concerns and sensibilities of the individual who discovers a new theory.” Kuhn describes this as the way “considerations relevant to the context of discovery [become] relevant to the context of justification” (360–61).

While Kuhn’s claim that the judgments of individual scientists guide theory choice should remind us of how Peirce, Turing, and Dreyfus each in their different ways showed how a set of programmed instructions does not capture what it is that a person does when he abducts hypotheses, that claim becomes more explicitly a matter of thinking through the relation between the sciences and the humanities in Kuhn’s 1989 essay “The Natural and the Human Sciences.” Here, Kuhn is careful to qualify his point about the way the context of discovery carries over to the context of justification, saying of the natural sciences that “though they may require . . . a hermeneutic base, [they] are not themselves hermeneutic enterprises” (222).<sup>10</sup> Conversely, “[v]ery little of what goes on” in the human sciences “resembles . . . research in the natural sciences,” since, in

the former, “new and deeper interpretations are the recognized object of the game” (222). So, for Kuhn, while in the natural sciences, the temperament and preferences of the individual person may come to play a role in theory choice (and shift) that doesn’t alter the basic distinction between disciplines that sometimes involve an interpretive element and those whose *raison d’être* is interpretation. Put differently, the humanities are not progressive disciplines in the sense that some new finding or discovery is going to make obsolete a prior interpretation or argument (comparable to, say, the way the discovery of oxygen, while an immensely complex matter of intellectual history, nevertheless led chemists simply to drop any references to “dephlogisticated air”).<sup>11</sup> If the problems and questions of the humanities invite newer and deeper interpretations, no amount of research is going to exhaust or dissolve such questions.

In a recent address at the National Humanities Center on “The State and Stakes of Literary Study,” Toril Moi described how graduate students in the humanities initiated into a *de rigueur* “hermeneutics of suspicion” have been led to the practice of debunking of artworks and so turning literary criticism into “exercises in fault finding.” As a corrective to this, she offers a conception of literary studies as an effort “to find the right terms in which to justify, defend and praise works of literature or other arts that merit our sustained attention.” *Praise*—a word Moi thinks too often calls to mind “thoughtless gushing rather than serious analysis”—is to be part of work in the humanities committed to understanding the “conditions of possibility for aesthetic utterance” and so of finding “serious intellectual and philosophical terms in which to convey to others why they should care” about particular artworks in place of what she describes as “dry as dust descriptions of some feature of the world *around* a work.”

If Moi is right that graduate training in literary studies tends to default to a predictable hermeneutics of suspicion—and thus, I would add, to a dispiriting uniformity, coterie-like groupthink and simple conformism—this perception is not unrelated to the scenario I have described here. The zeal to establish literary studies as a form of research and an attendant use of *discourse* understood as a general theoretic object in relation to which criticism is imagined as a science of discourse analysis are behind the idea that what is needed in response to the so-called crisis in literary studies is a theory of practice. I have been urging that we rather think of the discursive as that which makes possible the way persons intend things, take things to be thus-and-so, and project words into new contexts: for example, just those abilities, as we learned from Peirce, Dreyfus, Wittgenstein, and Cavell, that cannot be reduced to the operation of a set of rules. This alternative account becomes especially crucial for understanding what goes on in our responses to works of art. Instead of the

scientism explaining literature in accord with metrics borrowed from the sciences (authors are functions; artworks operate within discourse; text is a methodological field, and so on), we ought to treat artworks as intentionally structured unities that invite—indeed, offer dramatically heightened instances of—all the motley modes of response I have been calling the discursive. This approach would be something very like what Moi calls a search for the “right terms in which to justify, defend and praise works of literature or other arts that merit our sustained attention.” Indeed, what you study *is* the “attention,” just those acts of noticing and accounting, of argument and explication which seek to get something right about the idiosyncratic particulars of individual works. None of this is reducible to a general object or a formalizable theory, since art doesn’t make progress in the way the natural sciences make progress (there is no equivalent of “dephlogisticated air” in the history of artworks). Starting to think of what we do in literary studies as stemming from this sense of *discourse* rather than as something that does not belong to us might prevent the word from becoming at best a shibboleth and at worst a form of dogma.<sup>12</sup>

### Notes

1. For a thorough account of these vicissitudes, see Gerald Graff, *Professing Literature: An Institutional History* (1987).
2. See, for example, Paul Gross and Norman Levitt, *Higher Superstition: The Academic Left and its Quarrels with Science* (1994).
3. Recent tendencies in literary studies succumb in various ways to the allure of science as a model for grounding its practice in something respectably research-like; from so-called literary Darwinism, to scattered appeals to neuroscience, to the appropriation of statistics and statistical analysis for literary study, to what has come to be treated as the research program of “digital humanities.” For a compelling treatment of these issues in relation to literary Darwinism, see Jonathan Kramnick, “Against Literary Darwinism,” *Critical Inquiry* 37.2 (2011): 315–47. For a discussion of the current vogue for bringing statistics into the study of literature, see Williams, “The Statistical Turn in Literary Studies,” *Chronicle of Higher Education* 2 Jan. 2011: B14–15.
4. To be against the idea that literary studies is in need of a research program is by no means to be against the idea that literary studies involves scrupulous scholarship. Nor is my argument meant as an attack on colloquial uses of the word: for example, that an author is researching her novel set in Victorian London or that a parent concerned about bullying in her son’s school is doing research about the topic, and so on. By “research,” I mean a method theorized specifically in relation to clearly defined objects, with the aim of “producing new knowledge through metrics of evaluation emulating those of the sciences.”

5. For a more detailed account of Pierce's writings on logical machines see my *Experience and Experimental Writing* (2013), 46–49.

6. Turing's proposed criterion for what counts as artificial intelligence is an indistinguishability test in which a machine able convincingly to simulate a person's answers to a series of written questions can be called "intelligent." See Turing, "Computing Machinery and Intelligence," *Mind* 59.236 (1950): 433–60 and his lecture to the London Mathematical Society, 20 Feb. 1947 (in "Lecture on the Automatic Computing Engine." *The Essential Turing*, ed. Copeland [2004], 362–94).

7. While Turing is certainly an important precursor for work in AI research and computer science, he did not, as Copeland rightly notes, "work on anything than can be described as a AI program [rather] it is as the founder of the *philosophy* of AI that Turing takes his rightful place in the history of the subject" (Copeland, *Artificial Intelligence: A Philosophical Introduction* [1993], 253n30). AI research proper is thought to begin with the 1956 conference, the Dartmouth Summer Research Project on Artificial Intelligence. On the significance of the Dartmouth conference for initiating the AI research program, see Copeland, *Artificial Intelligence*, 8–9; 252–253n25–30.

8. The relevant passage in Wittgenstein is: "A rule stands there like a sign post.—Does the signpost leave no doubt open about the way I have to go? Does it show which direction I am to take when I have passed it; whether along the road or the footpath or cross-country? But where is it said which way I am to follow it; whether in the direction of its finger or in the opposite one?" (*Philosophical Investigations* [1951; 2009], 39). Wittgenstein is here building on earlier remarks from his 1933–1934 Cambridge lectures (collected in *The Blue and Brown Books: Preliminary Studies for the "Philosophical Investigations"* [1958]): "We give someone an order to walk in a certain direction by pointing or by drawing an arrow which points in the direction [but] couldn't such an order be interpreted to mean that the man who gets it is to walk in the direction opposite of that of the arrow?" (33).

9. Dreyfus remained resolute in his convictions about the limits of AI research, from the early strong AI models up to the more recent work on neural networks and connectionism. For an account of how even these more sophisticated connectionist models—which simulate learning through "feed-forward neural networks"—still "look hopelessly stupid" when compared with what persons take to be obvious for "getting around in the human world," see his "Intelligence Without Representation," Cognitive Sciences Initiative, University of Houston, 1998, web.

10. Kuhn does not foreclose the possibility that certain disciplines in the human sciences may at some point arrive at their own versions of "normal puzzle-solving" science, offering economics and psychology as examples (see "The Natural and the Human Sciences," 223).

11. The example is from Kuhn, *The Structure of Scientific Revolutions*, 94–98.

12. For help thinking through ideas central to the arguments in this essay, I am indebted to Rob Chodat, Kerry Larson, Nick Gaskill, Rebecca Newberger Goldstein, Chris Wood, and Dora Zhang.

### Works Cited

- Anderson, Douglas. "Peirce and Pragmatism: American Connections." *The Oxford Handbook of American Philosophy*. Ed. Cheryl Misak. New York: Oxford UP, 2010. 38–59.
- Barthes, Roland. "The Death of the Author." *Norton Anthology of Criticism and Theory*. Ed. Vincent Leach. New York: W.W. Norton, 2010. 1466–69.
- . "De l'Oeuvre au texte." *Roland Barthes Oeuvre completes*. Vol. 3. Paris: Editions du Seuil, 2002. 908–16.
- . "From Work to Text." *Norton Anthology of Criticism and Theory*. Ed. Vincent Leach. New York: W.W. Norton, 2010. 1470–75.
- Brandt, H. C. G. "How Far Should Our Teaching and Text-Books Have a Scientific Basis." *Transactions of the Modern Language Association of America* 1 (1884): 57–68.
- Cavell, Stanley. *Must We Mean What We Say?* Cambridge: Cambridge UP, 1969.
- Copeland, Jack. Headnote. "Systems of Logic Based on Ordinals (1938), Including Excerpts from Turing's correspondence, 1936–1938." *The Essential Turing*. Ed. Copeland. New York: Oxford UP, 2004. 125–45.
- Culler, Jonathan. "In Need of a Name? A Response to Geoffrey Harpham." *New Literary History* 36.1 (2005): 37–42.
- Dreyfus, Hubert. *What Computers Still Can't Do*. 1972. Cambridge: MIT P, 1992.
- Dyson, George. *Turing's Cathedral: The Origins of the Digital Universe*. New York: Pantheon, 2012.
- Foucault Michel. *The Archaeology of Knowledge*. 1969. New York: Vintage, 1982.
- . *L'ordre du discours*. Paris: Gallimard, 1971. Grimstad. "Experience and Experimental Writing." OUP, (2013).
- Harpham, Geoffrey Galt. "Beneath and Beyond the 'Crisis in the Humanities'." *New Literary History* 36.1 (2005): 21–36.
- Hodges, Andrew. *Alan Turing: The Enigma*. Princeton: Princeton UP, 1983.
- Hoy, David. "Foucault's Slalom." *London Review of Books* 4 Nov (1982): 18–20.
- Ketner, Kenneth Laine. "The Early History of Computer Design: Charles Sanders Peirce and Marquand's Logical Machines." *Princeton University Library Chronicle* 45.3 (1984): 187–224.
- Knapp, Jeffrey and Walter Benn Michaels. "Against Theory." *Critical Inquiry* 8.4 (1982): 723–42.
- Kuhn, Thomas. "The Natural and the Human Sciences." *The Road since Structure*. Chicago: Chicago UP, 2000. 216–43.



- . “Objectivity, Value Judgment, Theory Choice.” Machette Lecture, Furman University, 30 Nov. 1973. *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: U of Chicago P, 1977. 320–39.
- . *The Structure of Scientific Revolutions*. Chicago: U of Chicago P, 1962.
- Menand, Louis. *The Marketplace of Ideas: Reform and Resistance in the American University*. New York: W.W. Norton, 2010.
- Misak, Cheryl. “Charles Sanders Peirce (1839–1914).” *Cambridge Companion to Peirce*. Ed. Misak. Cambridge: Cambridge UP, 2004. 1–26.
- Moi, Toril. “The State and Stakes of Literary Study.” National Humanities Center. 20 Mar. 2010. “Toril Moi on ‘The State and Stakes of Literary Study.’” *Vimeo*. Vimeo, 2010. Web.
- Palermo, Charles. “Introduction: Intention and Interpretation.” *Nonsite 6* (2 Jul. 2012). Web.
- Peirce, Charles. “Logical Machine.” “Machine.” Def. 8. *The Century Dictionary and Cyclopedia*. Vol. 5. Ed. William D. Whitney. New York: Century Co., 1906. 3560.
- . “Logical Machines.” *American Journal of Psychology* 1 (1887): 165–70.
- . “Pragmatism as the Logic of Abduction.” *Essential Peirce: Selected Philosophical Writings*. Vol. 2. Ed. The Peirce Edition Project. Indianapolis: Indiana UP, 1998. 226–37.
- Sundquist, Eric. “The Humanities and the National Interest.” *American Literary History* 24.3 (2012): 590–607.
- Turing, Alan. “On Computable Numbers with an Application to the Entscheidungsproblem.” 12 Nov. 1936. *Proceedings of the London Mathematical Society* 2nd ser. 42 (1937): 230–65.
- Williams, Jeffrey. “The Little Magazine and the Theory Journal: A Response to Evan Kindley’s ‘Big Criticism.’” *Critical Inquiry* 39.2 (2013): 402–11.
- . “The Rise of the Theory Journal.” *New Literary History* 40.4 (2009): 683–702.