

# The Naturalist Turn in Literary and Art Study: Two-Way Interdisciplinarity and the Sciences of Mind

Patricia Kolaiti<sup>1\*</sup>

<sup>1</sup> Department of English Language and Literature, New York College, Athens, Greece

\* Patricia Kolaiti, E-mail: patriciakolaiti@hotmail.com

Received: April 27, 2017

Accepted: May 8, 2017

Online Published: May 15, 2017

doi:10.22158/selt.v5n2p277

URL: <http://dx.doi.org/10.22158/selt.v5n2p277>

*This work was supported by the AHRC [grant number 2004/111218], a UCL Graduate Scholarship, and the Balzan seminar on “Literature as an Object of Knowledge” (St. John’s College, Oxford).*

## **Abstract**

*In recent years, a burgeoning area of paradigm-revising scholarly investigation involves what could be referred to as a “Naturalist” or “Cognitive” turn in literary and art study, exploring the interface between theory in the arts and humanities and scientific theory of the type produced in disciplines belonging to the empirical and cognitive paradigm such as linguistics, cognitive science, philosophy of mind and cognitive neuropsychology. In this paper, I will discuss a range of theoretical, epistemic and methodological issues raised by such an interdisciplinary enterprise including the possibility of a genuine methodological merger with the cognitive paradigm, the plea for psychological realism, the extent to which the scientific method is compatible with the nature of literature and art as an investigative object and the need for genuine, two-way interdisciplinary practices in literary and art study. I will also briefly consider the role Relevance Theory might have to play in this interdisciplinary venture as both a pragmatic and epistemological framework.*

## **Keywords**

*“Naturalist” or “Cognitive” turn, interdisciplinarity, theory in the arts and humanities, the sciences of Mind, Relevance Theory*

## **1. Introduction**

After five decades of skepticism about the naturalist and cognitive paradigm, in the last few years there has been some evidence of a more systematic move towards a naturalist or cognitive turn in the Arts and Humanities, or at least, a systematic increase in empirical and cognitive awareness. At least two major interdisciplinary UK research projects systematically exploring the theoretical and methodological challenges raised by empirical disciplines—particularly domains such as linguistics,

cognitive science, philosophy of mind and cognitive neuropsychology—for contemporary literary scholarship and the philosophy of art have received funding from the Arts and Humanities Research Council (AHRC), UK (the Balzan project on “Literature as a Cognitive Object”, based at St. John’s College, Oxford, and the research project “Philosophical Aesthetics: the challenge from the sciences”, based at the University of Nottingham). At the same time, from 2012 the AHRC has been supporting the CogHumanities (Cognitive Futures in the Humanities) network, an interdisciplinary research network that brings together scholars from fields such as literature, linguistics, philosophy, etc., whose work informs, or is informed by, aspects of the cognitive, brain and behavioural sciences. These initiatives along with a growing body of cognitively and empirically aware literature (e.g., Anderson, 2015; Cave, 2016; Currie et al., 2014; Kieran, 2014; Lyne, 2014) can be seen as an integral part of a burgeoning area of paradigm-revising scholarly investigation that can be described as a “Naturalist” or “Cognitive” turn in literary and art study.

In this paper, I will discuss a range of key theoretical, epistemic and methodological issues raised by such an interdisciplinary enterprise, such as the possibility of a methodological merger of literary theory and the philosophy of art with the cognitive paradigm, the plea for psychological realism, the move from pre-theory to Theory, the extent to which the scientific method is compatible with the nature of literature and art as an investigative object and the need for genuine, two-way interdisciplinary practices in literary and art study. I will also briefly consider the role Relevance Theory might have to play in this interdisciplinary venture as both a pragmatic and epistemological framework.

## **2. Empirical and Cognitive Awareness in Literary and Art Study: Challenges from the Sciences**

To attempt to articulate *theory* within literary and arts study in the modern day is in a sense to plead for interdisciplinarity. Literature/art is not an autonomous object; and indeed, the literary/art event cannot be addressed by one discipline alone without being seriously diminished. It follows that any theoretical domain that hopes to take on literature/art as the global fact it is should not be anything less than an inter-discipline.

A century ago, literary scholars and philosophers of art would generally acknowledge the non-autonomous nature of the literary and art phenomenon, but at the same time, they could not be accused of theoretical misconduct for not using interdisciplinary explanatory tools, simply because there weren’t any. Today, there is no excuse. Moreover, the advances in the cognitive paradigm over the last thirty years, and the centrality that the concept of mind can be shown to have in the literary/art phenomenon, frame interdisciplinarity as a venture that calls for interaction primarily with cognitive domains. Although cognitive approaches to literary criticism and philosophical aesthetics have been sporadically pursued for some twenty or thirty years now, and words such as “cognitive” and “mind” are increasingly used in academic publications and titles of conferences or workshops in the last four to five years, the cognitive revolution still remains news to most mainstream literary and arts departments (for similar suggestions see Cave, 2011).

The story is old and familiar. Skepticism about the cognitive paradigm itself has very little to do with the specific proposals of cognitivism *per se*, and more with its departures from and implications for entrenched ideas about society, culture, religion, and the human being itself. Kuhn's (1996, p. 23) assertion that "paradigms gain their status because they are more successful than their competitors in solving a few problems that the group of practitioners has come to recognise as acute" comes across more as wishful thinking than as a realistic depiction of what leads to fluctuations in the fortunes of theories in the actual scientific world. Chomsky (1976, pp. 123-134) presented a more realistic picture by showing how ideological prejudice rather than reason was responsible for the commercial success of the more romantic Empiricism over the more scientifically sound and far better evidenced Nativism. Literary and art study are a paradigm example of the discipline-shaping effects of what Chomsky calls "ideological prejudice"; empiricist and behaviourist doctrines have taken deep root in traditional literary and art-philosophical scholarship, allowing socio-cultural approaches to almost monopolise the current perspective on literary and art phenomena, while disallowing cognitivist attempts to introduce considerations about, say, "mind", "agency" or "innateness" into the theoretical picture. At the same time, the mistaken idea that cognitivism stands in a competitive, rather than complementary, relation to the socio-cultural perspective, leading to ahistoricism and downgrading of the historically and culturally specific, further reinforces the skepticism of literary scholars about the cognitive paradigm.

Revising and updating literary-theoretical and art-philosophical discourse through an interdisciplinary merger with cognitive and empirical domains can be interpreted in a variety of ways. The most common view of the form such an interdisciplinary merger might take involves an increasing *awareness* of developments in the psychological sciences and cognitive disciplines (Currie et al., 2014). The plea for literary/art scholarship—and more generally, for scholarship in the Arts and Humanities, as the issues I raise here are no less relevant to disciplines such as, say, anthropology or cultural studies—to pay attention to such developments has been floating around for two to three decades now, and it is precisely this type of interdisciplinary merger that the quote from Richardson (1998) touches on. The radical advances from the 80's to the present day in fields such as cognitive science and psychology, cognitive anthropology, theoretical and experimental linguistics, philosophy of mind, philosophy of action, evolutionary and developmental psychology, psychology of perception and cognitive neuroscience, not to mention their implications for revising existing literary and art-theoretical issues and generating new ones, makes it unreasonable for literary and art study to continue working in isolation.

There is also a genuine question about whether there can be adequate theorizing in literary and art-study at all in the absence of any interaction with parallel developments in empirical and cognitive domains. Let me linger on this latter question a bit.

Reader-response criticism in the early 70's and in particular Roland Barthes' *The Death of the Author* (1967), which was particularly influential in literary study, focused on the reception end of the literary event and the reader's role in the "re-creation" of literary texts. In this seminal work, Barthes (1967, pp.

2-3; 1977, pp. 142-143) claims that “it is language which speaks not the author” and that “to write is to reach that point where language alone acts, ‘performs’, and not ‘oneself’”, discredits the role of authorial intentions in the process of both literary interpretation and production, while reference to the “chance” nature of the author’s/artist’s activity further discredits the reality of authorial intentions themselves. To what extent do Barthes’ assumptions correspond to what we have come to know in the last twenty years about intentionality, the recognition of others as intentional agents and the role of intentions in linguistic and non-linguistic communication?

Empirical work in cognitive science and developmental psychology suggests that the perception of others as intentional agents is fundamental to human experience and a cornerstone of social cognition. “A person’s bodily movements”, Woodward (2009, p. 53) suggests, are physically concrete. The goals and states of attention they embody are entirely abstract, yet real to human observers. We see others’ actions not as raw physical movements but, rather, as movements organized by intentional relations between agents and their goals and objects of attention. The ToM (Theory of Mind) capacity of attributing mental states to others, including intentional states, has its roots in early infancy and a rapidly growing body of research shows that much of what an infant learns in the first years of life depends on understanding others’ intentions (Tomasello et al., 2005; Tomasello & Moll, 2010). The earliest evidence that infants perceive others’ actions as intentional involves actions directed at concrete goals, such as reaching, and occurs as early as 5 months (Woodward, 1998, 2005, 2009). 5 month-old infants show selective attention to goals (e.g., they show increased looking at goal-oriented actions and do not respond in this way if the entity is not readily identified as an agent). There is also evidence that preverbal infants can identify emotional reactions that are incongruent with goal outcomes (Skerry & Spelke, 2014), that 18-month-old children are able to distinguish between inanimate objects (e.g. mechanical devices) and animate agents as well as between intended and unintended acts, and that they tend to reproduce the acts they regard as intended ones (Meltzoff, 1995, 2010; Hofer et al., 2005). The recognition of intentions seems to play a critical role in language acquisition too; there is evidence, for instance (e.g., Baldwin & Moses, 2001), that in learning new words, young children do not simply associate the words they hear with the things they happen to be seeing at the time, but follow the gaze and track the intentions of the person who uttered the word. Nowhere does the importance of mental state attribution for human functioning become more evident than in individuals with autism: autism has been linked to an impairment in ToM and is regarded as a kind of “mind-blindness” which involves difficulty in attributing mental states to others, including intentions (Baron-Cohen, 1995).

While cognitive science and developmental psychology provide evidence about the centrality of intention recognition to both individual and social cognition, empirical work in pragmatics since the “70s on how hearers bridge the gap between encoded/linguistic meaning and communicated meaning—or more simply, between what words mean in their own right and what a speaker means in using these words—suggests that the recognition of speakers’ intentions has a fundamental role to play in the interpretation of utterances. From Grice’s (1989) early work on how hearers infer a speaker’s

conversational implicatures to more fully developed theories of communication such as Relevance-theoretic pragmatics (Sperber & Wilson, 1995; Wilson & Sperber, 2012), human communication emerges as heavily inferential, with the recognition of speaker intentions occupying a central place.

In this light, Barthes' attempt to present the author as an unintentional agent, deny the reality of authorial intentions and discredit their role in the process of literary interpretation clearly lacks what later on I will call *psychological realism*: Barthes' account is incompatible with the accumulating empirical evidence about the foundational place of intention recognition in both the human cognitive make-up and the mechanisms that make linguistic—and consequently literary—communication possible, and is therefore incompatible with the psychological reality of how the human mind actually works. Any attempt to articulate theory within literary and art study without incorporating pertinent empirical findings and theories from the human sciences runs the risk of yielding theoretically inadequate and psychologically unrealistic accounts.

### **3. From Pre-Theory to Theory: Interdisciplinarity and Methodological Merger**

For many decades, in the Arts and Humanities we have been describing loosely as “interdisciplinary” intellectual outputs that are better described as multi-disciplinary (for similar views see Richardson, 2010, p. x; Heintz, Origgi, & Sperber, 2004) outputs that indeed involve selective reading across more than one discipline, but more often than not do not require any essential interaction or merging between the methods and reasoning styles of the disciplines themselves.

For the type of interdisciplinary merger discussed here to be genuine and intellectually informed, a condition of methodological merger must also be fulfilled: genuine interdisciplinary interaction is not simply a matter of sharing topics, but also, and more crucially, of merging methods and mind-sets. The literary/art-theoretical world needs to do more than simply delve into empirical topics; what it also needs is to start revising its established investigative and paradigm-specific practices. Understood in this way, genuine interdisciplinarity between literary/art—theoretical and empirical/cognitive domains calls not only for extensions to the theories involved but also for a broader paradigm shift. In many respects, this paradigm shift can be seen as involving what could be described as an initial move from pre-theory to Theory.

So far, I have talked about theory in an undifferentiated sense. Now, I would like to talk briefly about literary and art Theory in a robust sense of the term, thereby contrasting “Theory” with “pre-theory”.

Let us take “Theory” to mean “theory compatible with the scientific method”, where by “scientific method” I mean the method of the Natural Sciences. Roughly speaking, a theory compatible with the scientific method aims at testable claims and empirical hypotheses, that is claims and hypotheses amenable to scientific investigation, and may be described as a complex articulated body of thought geared to answering questions or explaining regularities in a given domain. A Theory purports to make true claims, and to make a true claim you have to make a prediction about the (real) world. Truth is not

a property either of objects or of states of affairs in the world, but of (possible or actual) thoughts. It involves a particular type of correspondence between thoughts entertained as true descriptions of the world and the states of affairs they describe. A Theory is tested by testing its implications—Van der Henst and Sperber’s “Testing the cognitive and communicative principles of relevance” (2004) discusses in great detail how theories are tested by testing their implications, and is confirmed to the extent that its implications are both testable and true:

If the Theory (T) is correct, it should follow that P.

Here’s some evidence that P.

P supports my Theory (T) and provides some evidence that T’s claims are correct.

Or, by contrast:

If the Theory (T) is correct, it should follow that P.

Here’s some evidence that P is not the case.

P supports a Theory (T2) that is incompatible with my Theory (T) and provides some evidence that T’s claims are incorrect.

And, finally, Theories compatible with the scientific method are arrived at by non-demonstrative inference: the truth of their implications is typically not guaranteed by the truth of the premises, but merely made more probable. In this light, a Theory can be seen as a complex, non-demonstratively evidenced body of thought: a body of articulated assertions subject to constant modification and revision in the light of new evidence. Synchronically and diachronically, a Theory is constantly assessed and modified in the direction of increasing effectiveness and simplicity, i.e., in the direction of maximising the number of true implications (and other cognitive effects) and achieving these effects more elegantly.

In the cognitive sciences, Theory is also expected to have *explanatory value*. The notion of “explanatory” value can be contrasted with “descriptive” value: the aim of Theory is not simply to describe objects, states of affairs in the world or regularities in a given domain but also to provide a systematic and articulated body of explanations that account for how these regularities or states of affairs come about. In this respect, the process of explanation is tightly interwoven with the investigation of causation, whilst causal considerations are entirely irrelevant to a descriptive project. In domains such as the study of human natural language, a degree of tension may arise between the goals of description and explanation (e.g., see Chomsky, 2000).

At the same time, to say that a Theoretical account in a domain that involves human cognition is “explanatory” is to imply that it is *psychologically real*: for a Theory to be explanatory it must provide a psychologically realistic account of the regularities or states of affairs it aims to explain in the sense that the mechanisms it describes must correspond to those the human mind/brain actually deploys.

Finally, the articulation of Theory compatible with the scientific method goes hand-in-hand with a characteristic mind-set or reasoning style. Making rough generalizations, we could say that Theory tries to avoid ambiguity and vagueness. This is not to deny, of course, that in developing a theory of any

type, theorists start with initial speculative conceptions and vague thoughts. Without this vague pre-theoretical stage, theory would never get off the ground. This is one of the two reasons I see why the articulation of Theory does not in any way suggest a total elimination of pre-theoretical discourse. The other concerns the empirically attested fact that scientific claims and hypotheses are sometimes expressed in a vague and pre-theoretical form, and it may be very long after their initial articulation that the Theorist herself or some other Theorist manages to express them in empirically tractable Theoretical form. Theory and pre-theory do not necessarily stand in a competing relationship and, to my mind, an adequate understanding of the nature of Theory also involves a need to grasp the parallel importance of pre-theoretical thinking for human intellectual and scientific development. Yet the aim of Theory is to achieve the greatest possible clarity in expression by progressively replacing ambiguous and vague pre-theoretical claims with unambiguous and testable ones. Theory favours assertions that are amenable to strong confirmation or disconfirmation, aims at subtle distinctions, requires argument often further supported by empirical and experimental evidence, tends to focus on narrow and specialized problems and values detail at both the level of explication/analysis and that of investigation. A paradigm-shift in the direction of Theory in literary-theoretical and art-philosophical practice is not a straightforward or hassle-free enterprise. A subtle balance needs to be maintained between preserving the special character of established literary and art-theoretical investigative routines, while at the same time putting into operation those new paradigm-changing practices that could—at least partly—revise literary and art-philosophical discourse in the direction of Theory and create the necessary space for constructive interdisciplinary interaction with a broad range of pertinent empirical, quasi-empirical and cognitive domains. Earlier in my discussion, for instance, I tried to underline the importance of pre-theoretical thinking for the investigative practices of even robustly empirical fields.

Pre-theoretical thinking involving half-understood beliefs of a certain kind is integral to what I called the “special character” of literary/art-theoretical intellectual endeavours. A notion of *intuitive half-understood beliefs* may allow some insight into why that is. In discussing belief-holding, Sperber (1997, p. 67) introduces the term *reflective beliefs*, i.e., beliefs which involve partially understood thoughts or thoughts containing a partially understood constituent, putting the weight on the credibility and authority of the source of the belief: “Reasons to hold reflective beliefs are provided by other beliefs that describe the source of the reflective belief as reliable, or that provide explicit arguments in favour of the reflective belief” (Note 1).

It could be argued, however, there should also exist intuitive reasons for holding half-understood beliefs: these should be seen as independent of the credibility and/or authority of the source and dependent only on (pre-rational) intuitions about the relative importance of the implications of the thought. Such beliefs are half-understood but not strictly speaking “reflective”; they are not believed metarepresentationally—because a credible source says so—but intuitively -because of pre-rational pieces of introspective evidence. This introspective evidence we might or might not be able to rationalize and explicate but this does not affect the relative importance of the implications of the

thought. It is this latter type of half-understood beliefs that I would like to call *intuitive half-understood beliefs*, taking them to be sharply different from reflective half-understood beliefs. To come to the point, intuitive half-understood beliefs are integral to literary and art-theoretical thinking and although from a more robustly empirical viewpoint they could indeed be seen as mere precursors to adequately Theoretical assertions, the relative importance of their implications cannot be discounted.

Alongside the issue of how to revise paradigmatic practices while preserving the (healthy) aspects of the established character of literary and art-theorising, it also seems that, at present, a good proportion of the literary and art-philosophical world is not convinced that Theory, and more generally the *scientific method*, is appropriate to the nature of literature and art as an investigative object in the first place.

Despite the literary and art world's skepticism about the methods of the natural sciences, the fact remains that not a single compelling reason has ever been proposed for why literature and art might not be amenable to scientific investigation. Science aims at understanding; yet not all types of understanding are compatible with science. Scientific understanding is concerned with causation and explanation. And from this vantage point, many more aspects of literature and art than is standardly supposed can be shown to be relevant to scientific enquiry.

To give just one example, the institutional and historical dimension of art, often grouped together with those aspects of the literary and art phenomenon that are not relevant to scientific investigation, might inform and be informed by cognitive research on the role of *mental prototypes/prototypical features* and *prototypicality judgements* (e.g., Armstrong, Gleitman, L. R., & Gleitman, H., 1983; Barsalou, 1987, 1992; Malt & Johnson, 1992; Landau, 1994) in the process of categorization, i.e., in classifying an object as belonging to a certain category. Established institutional practices may be grouped together with a diverse range of factors affecting an object's probabilities of being recognised and categorised as such and such (e.g., as an ARTWORK) in a given spatiotemporal context, while the historicity of art may be considered in relation to how historically stabilised chains of prototypical resemblances facilitate or even enable the recognition of certain objects—even highly atypical ones—as such and such in the long process of human cultural transmission.

The most persistent reservation, however, about the ability of science to solve problems and answer questions relating to art seems to me to have its roots in what we could call “the problem from conceptual analysis”. “Conceptual analysis aims to find informative, necessary, and sufficient conditions for the application of a concept, but the problems that people encountered with *art* suggested to some that in this case at least the project is hopeless rather than merely difficult” (Currie et al., 2014, p. 5). Philosophers have repeatedly pointed out the failure of theory to subject the concept LITERATURE and ART to adequate analysis in the sense of finding a “definition” for it. Back in 1956, Weitz's attempt to explain the grounds of this failure is particularly important:

Is aesthetic theory, in the sense of a true definition or set of necessary and sufficient properties of art, possible? If nothing else does, the history of aesthetics itself should give one enormous pause here. For,



in spite of the many theories, we seem no nearer our goal today than we were in Plato's time. (...) I want to show that the inadequacies of the theories (...) reside (...) in a fundamental misconception of art. Aesthetic theory—all of it—is wrong in principle in thinking that a correct theory is possible because it radically misconstrues the logic of the concept of art. Its main contention that “art” is amenable to real or any kind of true definition is false. Its attempt to discover the necessary and sufficient properties of art is logically misbegotten for the very simple reason that such a set and, consequently, such a formula about it, is never forthcoming. Art, as the logic of the concept shows, has no set of necessary and sufficient properties, hence a theory of it is logically impossible and not merely factually difficult. Aesthetic theory tries to define what cannot be defined in its requisite sense (Weitz, 1956, pp. 27-28).

Further on in his discussion, Weitz (1956, p. 31) explains that theory fails to find a definition for the concept ART because ART is an “open” concept: as a result of the artist's tendency to innovate, the concept ART can be repeatedly applied to new instances which alter the set of necessary and sufficient conditions and throw our prior definitions into disarray. He therefore concludes that there can be no defining conditions for arthood.

I take the problem from conceptual analysis to be closely related to the reservations of literary and art scholars about the extent to which art may be amenable to scientific investigation. The quote from Weitz, for instance, explicitly associates the difficulty of finding a definition for the concept ART with the appropriateness of theory for the aims of philosophical aesthetics. More generally, the problem from conceptual analysis may be seen as having led to a conviction that art is too unstable or too mysterious an object to be subjected to empirical analysis of any sort. I would therefore like to take a moment to show that a) the failure to find a definition of ART is not really evidence that art is not a suitable investigative object for Theory, and b) the problem from conceptual analysis itself is fundamentally flawed.

For the best part of the 20th century, it seemed that the so-called *classical* or *decompositional* or *definitional theory of concepts* (Kripke, 1972) was the only theory that could explain both the role of concepts as word meanings and their role in cognition and categorisation. The definitional theory of concepts had its roots in early lexical semantics—which was itself inspired by the empiricist idea that complex concepts are made by gluing together simple concepts—and claimed that concepts (and consequently word meanings) must have the form of a definition, i.e., that they decompose into a set of necessary and sufficient conditions for an item to fall under the concept. On this approach, the concept BACHELOR would decompose into the more basic concepts HUMAN & ADULT & MALE & NOT-MARRIED. This definition would function as the meaning of the word “bachelor”, and provide a set of necessary and sufficient conditions enabling us to categorise bachelors and recognise them as distinct from other things.

Although there is something strongly appealing about this approach—indeed, in everyday life or in, say, learning a second language, we effortlessly exchange information about the meanings of words using definitions, and this is also how dictionaries work—the decompositional theory of concepts has been

increasingly criticized in recent years. For one thing, most concepts just aren't like BACHELOR. For most concepts say, TRUTH, or TIGER—it seems impossible to find necessary and sufficient conditions. For another, Jerry Fodor's (1975) book *The Language of Thought*, and Fodor, J. A., Fodor J. D. and Garrett's (1975) article "The psychological unreality of semantic representations" attacked the definitional account and provided compelling evidence that the view of concepts as definitions is inadequate in several respects: as a theory of word meaning, as a theory of categorisation and as a theory of concept acquisition. In the last 25 years or so, Fodor has argued at length that word meanings—even of words like "bachelor"—are not definitions, cognitive psychology has brought to light accumulating evidence that we do not categorise objects by checking them against a definition and research in concept acquisition offers a growing body of evidence that children are born with a considerable amount of conceptual knowledge that is not acquired through the senses. Work in these areas now strongly points in the direction of *conceptual atomism*, suggesting that the meanings of most words are not definitions but simple, unanalysable or *atomic* concepts: the meaning of the word "bachelor" is the simple atomic concept BACHELOR (see Lawrence & Margolis, 2002, pp. 8-24, pp. 52-59 for criticisms of the classical approach; and Lawrence & Margolis, 2002, pp. 59-71; Margolis, 1998 for a summary of Fodor's views).

Theory has failed to decompose ART into a definition just as it has failed to decompose any other concept. This failure, however, is not down to literature and art's mysterious, unstable, "open" or other nature, but to the fact that the very enterprise of conceptual analysis and the decompositional or definitional theory of concepts it was based on is inadequate in the first place. The impossibility of decomposing ART into an informative and stable set of necessary and sufficient conditions is not evidence for the inadequacy of Theory or the scientific method to investigate literature and art but only further evidence for the inadequacy of the decompositional theory of concepts itself (Note 2).

So I am back to where I started: not a single compelling argument has ever been put forth to show that there is a mismatch between literature and art as an investigative object and the methods of the natural sciences.

#### **4. Methodological Merger and the Role of Relevance Theory**

In recent years, a few cognitively oriented literary and art scholars (Cave, 2016; Cave & Wilson, forthcoming) have turned their attention to advances in Pragmatics and begun considering whether Relevance Theory (Sperber & Wilson, 1995), or at least some of the theoretical concepts it has developed, might have a role to play in reviving and updating literary-theoretical discourse, and if so, what role this may be (Note 3).

*Relevance* is not only a theory of communication and cognition but also a significant epistemic step. Amongst other things, *Relevance* broadened the range of explanations available to the humanities, allowing phenomena to be accounted for in psychologically realistic terms as opposed to being merely described. It replaced pre-theoretical discourse with testable Theoretical principles, revived

long-forgotten questions in a surprisingly wide range of other disciplines and made pragmatics one of the most influential interdisciplinary ventures of our times. Drawing on the paradigm of the natural sciences, *Relevance* attempted a leap from pre-theory to Theory via a body of theoretical suggestions compatible with the scientific method. Still, what literary and art theory could gain from a collaboration with Relevance-theoretic pragmatics does not stop at purely epistemological rewards.

A common thread running through most recent work in poetics and stylistics, for instance, no matter how diverse the theoretical affiliations of the author, is a universal discontent with the simple, fixed, binary oppositions of structuralist linguistics and the so called “Bi-planar” or “Code model of Communication”, and a parallel desire to locate stylistic enquiry within pragmatic research (Leech, 1983; Fabb & Durant, 1987b, pp. 229-237; Kiparsky, 1987, p. 185; Fowler, 1996, pp. 199-200; Toolan, 1996, pp. 121-124). In addition to the text—which has almost monopolised the attention of stylistics in the last fifty years, readers and their *cognitive environment* must now also be taken into account. Relevance Theory has developed an elaborate notion of a cognitive environment (Sperber & Wilson, 1995, pp. 38-39) that accounts for inter-individual differences as well as sharing of experience in explanatory and empirically tractable terms. The notion of a cognitive environment is also fundamental to understanding and explaining the cognitive infrastructure of a broad range of phenomena integral to the literary/art event, from interpretive convergence and divergence both across individuals and across interpretive communities to the way socio-historical context may affect the reception and recognition of an artwork or literary text, to the mechanisms that make installation-related decisions crucial to the way an artwork may be aesthetically experienced and received.

Note also that there has never been, and it is hard to imagine how there could be, a literary and art theory without an underlying theory of communication and cognition. Although particularly influential in the area of Pragmatics, Relevance-theory provides not only one of the most fully developed theories of communication but also a sophisticated theory of human cognition. Based on the fundamental assumption that our mind has evolved in the direction of increasing *cognitive efficiency*, so that it tends to use its attentional and other processing resources in the most productive way, Relevance theory proposes a cognitive account (Sperber & Wilson, 1995, pp. 38-46, pp. 260-266) of how attention is allocated among the indefinite range of inputs that are perceptible or inferable in our physical and mental environment at any one time: functioning as an efficient information processing device, the mind will develop automatic heuristics for allocating attention to the inputs with the greatest expected *relevance*, i.e., to those inputs that are expected to achieve the greatest cognitive effect for the least possible processing effort. Human cognition is naturally geared to the maximisation of relevance. The Relevance-theoretic view of human cognition provides the basis for developing psychologically realistic accounts of the evolutionary adventure of literature and art and the centrality of the literary/art event in human social and cultural transmission. It also motivates a range of fascinating questions of a broad interdisciplinary scope concerning the kind of stimulus that a work of literature and art is and the types of effects that have made, and continue to make, artistic stimuli relevant to individual and

collective cognition.

Owing a lot to the ideas of Paul Grice, whose work on meaning was the first to give an important place to a speaker's intentions and the role the recognition of these intentions by an audience/hearer might play in interpretation, Relevance theory presents one of the most fully developed accounts of human ostensive inferential communication. Indeed, in the Relevance-theoretic framework, the study of linguistic communication and the way hearers bridge the gap between encoded and communicated meaning is seen as merely a special case of the broader phenomenon of inferring a communicator's meaning from ostensive stimuli: linguistic communication, just as much as non-linguistic communication, relies on the production of an ostensive—or more simply, communicative—stimulus whose function is to attract a hearer's attention and focus it on the communicator's intentions. In this way, Relevance theory offers perhaps the most wide-reaching theory of communication available and provides a refined explanatory model that allows literary and art-theorising to account for various aspects of the construction of meaning in literary and art reception, the complicated interplay between conceptual representations/propositional content and the elicitation of aesthetic experience in literature and art and, finally, the active engagement of the receiver in the construction of meaning and experience.

Relevance theory provides a wide-ranging and refined model of communication and a psychologically realistic account of human cognition that both literary and art theory can certainly make use of. Ultimately, given the exemplary interdisciplinary scope of Relevance theory, delving into the issues it discusses could provide the contemporary literary and art theorist with immediate insights into the issues and advances of the many disciplines with which Relevance theory interacts. Yet in a healthy and full-blown interdisciplinary dialectics, and always given the non-autonomous nature of literature and art, Relevance theory should also have various rewards to reap from an interaction with literary and art-theoretical domains. The exchanges should be expected to go both ways, which brings us to the issue to be discussed in the next and final section of this paper.

## **5. Two-Way Interdisciplinarity and Bi-Directionality of Effects**

So far I have looked at two different ways in which an interdisciplinary merger between literary/art study and cognitive domains can be conceived: one sees interdisciplinary merger as a matter of empirical and cognitive awareness in literary and art study; the other sees interdisciplinary merger as a case of methodological integration, expecting it to induce significant discipline-revising changes in the literary and art-philosophical paradigm. There is one further crucial sense in which a genuine and full-blown interdisciplinary merger with cognitive and empirical domains can be thought of: for a genuine interdisciplinary merger to take place, theorizing in literary and art study should be expected to have retroactive effects on theories produced in the full range of disciplines with which it interacts.

Interdisciplinarity, as I use the term here, involves a reciprocal, two-way or bi-directional relationship between two disciplines, such that theory and practice in the one discipline can, at least in principle,

have a direct bearing on theory and practice in the other, and conversely. It is hard to imagine how else it could be. Hard to imagine, for instance, that theorizing in the Philosophy of mind (on the issue of *mental architecture*, for example) could have major implications for Pragmatics by raising questions such as “How is Pragmatics located within the broader architecture of the mind?”, but theorizing in Pragmatics would not have retroactive effects on the Philosophy of mind. The very minute Pragmatics contemplates the “mental location” of our pragmatic mechanisms, theory in the Philosophy of mind is instantaneously affected. This, if nothing else, is a first-rate example of a genuine interdisciplinary relationship. To give another example, pragmatic research on the nature of communicator’s meaning should normally be expected to have direct implications for understanding how communicator’s meaning is conveyed in literary texts; at the same time, investigating the sort of communicator’s meaning conveyed in some literary works should be expected to lead retroactively to a radical enrichment of the notion of communicator’s meaning itself.

Lacking a proper theoretical apparatus of its own, literary and art study have so far been borrowing from anthropology, linguistics, psychology and the many other human-scientific domains that have entered the game of inter-blend in the last fifty years, but without necessarily considering whether it could or should give something back. Many theorists (e.g., Jackson, 2002, pp. 177-178) have seen this uni-directional game as the only realistic possibility for literary/art study, and dismissed the idea of two-way interdisciplinarity as unattainable, particularly when the other side of the inter-blend involves cognitive domains. Traditional theorists express skepticism about any transformations of the discipline that might result from adopting naturalistic investigative practices, and are concerned that, as soon as literary/art theorizing starts using investigative methods or addressing questions that will give it the “epistemological weight to affect theory” in other, particularly empirical, disciplines, the resulting investigative practice simply will not be literary/art theorizing any more.

Generally speaking, reservations about any kind of two-way interdisciplinary effects could be associated with the epistemological concern I discussed earlier in this analysis, i.e., the concern of literary and art scholars about how far scientific disciplinary practices or, more simply, the scientific method, are appropriate to the nature of literature and art as an investigative object. As I tried to show earlier, the assumption that the scientific method is incompatible with the nature of literature and art as an object of enquiry seems largely unsubstantiated, and if so, then the incorporation of naturalistic practices and the bi-directionality of interdisciplinary effects that may result from them poses no serious threat to the disciplinary integrity of literary and art theorizing.

The bi-directionality of a genuine, two-way interdisciplinary relation seems to raise two parallel questions. The first is how pertinent scientific enquiry can be of use to literary/art studies. My earlier discussion of empirical and cognitive awareness in literary and art study touches on precisely this issue. And indeed, a few steps seem to have been made in this direction as a number of contemporary literary and art paradigms have been drawing increasingly on scientific enquiry in one way or another. The second question arises only on the assumption that literary/art theorizing can retroactively affect theory

in other disciplines, and this may explain why, although equally important for any worthwhile notion of interdisciplinarity, it has not received the same amount of attention. This question reverses matters and asks how (theoretical) literary/art studies can be of use to pertinent scientific enquiry.

Particularly intricate, highly self-reflexive and probably unique as an action-process, the production of literature and art is relevant to cognitive science, philosophy of action, philosophy of mind, neuropsychology and theoretical linguistics.

The empirical fact that the making of literature/art is an internally and to a great extent *spontaneously caused* creative activity, raises significant interdisciplinary issues about human agency, intentionality and the nature of action, alongside genuine questions about inspiration, causation, consciousness and free will.

Creativity and the way it manifests itself in the making of literature and art is a point on which the study of artistic thinking could contribute to cognitive science, creativity-related research and the study of human practical reasoning processes, furthering our understanding of the nature of generic forms of creativity as well as of the interface or divergence between artistic thought and other (plausibly specialised) varieties of creative thinking such as scientific creativity.

At the same time, the subtle cognitive, perceptual and kinaesthetic abilities that seem associated with literary and artistic production make the empirical study of literary/artistic practice relevant for neuroscience, cognitive science and developmental psychology with a view to illuminating the kinaesthetic, perceptual and cognitive infrastructure of human creativity in general.

The role of imagination in the literary and art event renders artistic production and reception fruitful domains for exploring the nature of human imagining abilities, while the role of metarepresentation in both the production and reception of literature and art and the way literary and artistic response relies on it or disrupts it, make the study of the literary and art event particularly important for our grasp of the human metarepresentational capacity and the range of domains in which Theory of Mind came to be used in the long course of human evolution.

The crucial role of subconscious thinking in the creation of literature and art makes literary and artistic production particularly important for understanding the intricate workings of sub-attentive thought processes, their interplay with conscious thought, their contribution to practical reasoning and their role in the relation between language and thought.

The fact that artistic communication relies heavily on weak implications/implicatures and poetic effects makes the study of *weak* forms of communication in art potentially interesting for current debates in pragmatics and cognitive science, while investigating the type of communicator's meaning conveyed by some literary works could contribute to current pragmatic and cognitive research in areas such as linguistic underdeterminacy, *ad hoc* concept formation, cognitive plasticity, local v global relevance etc, leading to a radical enrichment of the notion of communicator's meaning itself.

The pleasurable and often intensely physical aspect of the characteristic type of experience elicited by artworks (aesthetic experience) suggests that an empirically-oriented study of the aesthetic is likely to

enrich our broader understanding of the human ability for pleasurable response from an evolutionary, neuroscientific and cognitive perspective. This, in a certain sense, is synonymous with an investigation of the relevance-yielding aspect of artistic stimuli; or, to put it more simply, an investigation of the properties that make artistic stimuli worthy of individual and collective attention, and are responsible for the success of literature/art as both a public, culturally situated entity and an occurrence within human individual consciousness. Such an investigation is likely to have rich implications for domains such as pragmatics, neuroscience, cognitive anthropology, cognitive cultural studies, evolutionary and cognitive psychology, and for notions such as cognitive efficiency, the nature of relevance and the types of effects that make a stimulus worth attending to.

These are just a few of the questions about which a genuine, two-way-interdisciplinary literary and art-theoretical discourse should have something to say. Literature and art is so characteristic an output of the human mind, that a “Cognitive” or “Naturalist” turn in literary and art study cannot be realized in any robust sense of the term unless literary and art-theoretical inquiry start systematically pursuing bi-directional effects on the full range of empirical and cognitive disciplines with which they interact.

## References

- Anderson, M. (2015). *The Renaissance Extended Mind*. UK: Palgrave Macmillan.
- Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What Some Concepts Might Not Be. *Cognition*, 13(3), 263-308. [https://doi.org/10.1016/0010-0277\(83\)90012-4](https://doi.org/10.1016/0010-0277(83)90012-4)
- Baldwin, D. A., & Moses, J. A. (2001). Links between social understanding and early word learning: Challenges to current accounts. *Social Development*, 10, 311-329. <https://doi.org/10.1111/1467-9507.00168>
- Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.
- Barsalou, L. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and Conceptual Development*. Cambridge: Cambridge University Press.
- Barsalou, L. (1992). Frames, concepts, and conceptual fields. In E. Kittay, & A. Lehrer (Eds.), *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization*. Hillsdale, NJ: Lawrence Erlbaum.
- Barthes, R. (1967). *The death of the author*. Retrieved from [http://www.tbook.constantvzw.org/wp-content/death\\_authorbarthes.pdf](http://www.tbook.constantvzw.org/wp-content/death_authorbarthes.pdf)
- Bond, E. J. (1975). The Essential Nature of Art. *American Philosophical Quarterly*, 12, 177-183.
- Cave, T. (2011). *Thinking with Literature* (talk delivered at the Balzan workshop on “Concepts”, St John’s College, Oxford).
- Chomsky, N. (1976). *Reflections on Language*. London: Temple Smith.
- Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*. Cambridge University Press.
- Currie, G., Kieran, M., Meskin, A., & Robson, J. (Eds.). (2014). *Aesthetics and the Sciences of Mind*.

- Oxford: Oxford University Press.
- Fabb, N., & Durant, A. (1987). New courses in the linguistics of writing. In N. Fabb et al. (Eds.), *The Linguistics of Writing; Arguments between Language and Literature*. Manchester University Press.
- Fodor, J. A. (1975). *The Language of Thought*. Cambridge, MA: Harvard University Press.
- Fodor, J. A., Fodor, J. D., & Garrett, M. (1975). The psychological unreality of semantic representations. *Linguistic Enquiry*, 6, 515-531.
- Fowler, R. (1996). Studying literature as language. In J. J. Weber (Ed.), *The Stylistics Reader*. London: Arnold.
- Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Hofer, T., Hauf, P., & Aschersleben, G. (2005). Infant's perception of goal-directed actions performed by a mechanical device. *Infant Behavior and Development*, 28, 466-480. <https://doi.org/10.1016/j.infbeh.2005.04.002>
- Jackson, T. E. (2002). Issues and problems in the blending of cognitive science, evolutionary psychology and literary study. *Poetics Today*, 23, 161-179. <https://doi.org/10.1215/03335372-23-1-161>
- Kieran, M. (2014). *Philosophical Aesthetics and the Sciences of Art, Royal Institute of Philosophy Supplement* (Vol. 75, pp. 203-230). Cambridge University Press.
- Kiparsky, P. (1987). On theory and interpretation. In N. Fabb et al. (Eds.), *The Linguistics of Writing; Arguments between Language and Literature*. Manchester University Press.
- Kripke, S. (1972). Naming and necessity. In D. Davidson, & G. Harman (Eds.), *Semantics of Natural Language*. Dordrecht: D. Reidel.
- Kuhn, T. S. (1996). *The Structure of Scientific Revolutions*. Chicago and London: The University of Chicago Press.
- Landau, B. (1994). Where's What and What's Where: The Language of Objects in Space. *Lingua*, 92, 259-296. [https://doi.org/10.1016/0024-3841\(94\)90344-1](https://doi.org/10.1016/0024-3841(94)90344-1)
- Laurence, S., & Margolis, E. (2002). Radical concept nativism. *Cognition*, 86, 25-55. [https://doi.org/10.1016/S0010-0277\(02\)00127-0](https://doi.org/10.1016/S0010-0277(02)00127-0)
- Leech, G. (1983). *Principles of Pragmatics*. London: Longman.
- Lyne, R. (2014). *Shakespeare, rhetoric and cognition*. Cambridge University Press.
- Malt, B. C., & Johnson, E. C. (1992). Do Artifact Concepts Have Cores? *Journal of Memory and Language*, 31(2), 195-217. [https://doi.org/10.1016/0749-596X\(92\)90011-L](https://doi.org/10.1016/0749-596X(92)90011-L)
- Margolis, E. (1998). How to acquire a concept. *Mind and Language*, 13, 347-369. <https://doi.org/10.1111/1468-0017.00081>
- Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, 31(5), 838-850. <https://doi.org/10.1037/0012-1649.31.5.838>
- Meltzoff, A. N. (2010). Bridging between action representation and infant theory of mind. In P. A.



- Frensch, & R. Schwarzer (Eds.), *Cognition and Neuropsychology* (Vol. 1). International perspectives on psychological science. New York: Psychology Press.
- Richardson, A. (1998). Brains, Minds, and Texts. *Review*, 20, 39-48.
- Richardson, A. (2010). *The Neural Sublime: Cognitive Theories and Romantic Texts*. Baltimore: The John Hopkins University Press.
- Skerry, A. E., & Spelke, E. S. (2014). Preverbal infants identify emotional reactions that are incongruent with goal outcomes. *Cognition*, 130, 204-216. <https://doi.org/10.1016/j.cognition.2013.11.002>
- Sperber, D. (1997). Intuitive and reflective beliefs. *Mind and Language*, 12, 67-83. <https://doi.org/10.1111/1468-0017.00036>
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition* (2nd ed.). Oxford: Blackwell.
- Tomasello, M., & Moll, H. (2010). The gap is social: Human shared intentionality and culture. In P. Kappeler, & J. Silk (Eds.), *Mind the gap: Tracing the origins of human universals*. Berlin: Springer.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675-735. <https://doi.org/10.1017/S0140525X05000129>
- Toolan, M. (1996). Stylistics and its discontents; or, getting off the Fish "hook". In J. J. Weber (Ed.), *The Stylistics Reader*. London: Arnold.
- Turner, M. (2002). The cognitive study of art, language and literature. *Poetics Today*, 23, 9-20. <https://doi.org/10.1215/03335372-23-1-9>
- Van der Henst, J. B., & Sperber, D. (2004). Testing the cognitive and communicative principles of relevance. In I. Noveck, & D. Sperber (Eds.), *Experimental Pragmatics*. Basingstoke: Palgrave-Macmillan.
- Weitz, M. (1956). The Role of Theory in Aesthetics. *The Journal of Aesthetics and Art Criticism*, 15, 27-35. <https://doi.org/10.2307/427491>
- Wilson, D., & Sperber, D. (2012). *Meaning and Relevance*. Cambridge University Press.
- Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, 69(1), 1-34. [https://doi.org/10.1016/S0010-0277\(98\)00058-4](https://doi.org/10.1016/S0010-0277(98)00058-4)
- Woodward, A. L. (2005). The infant origins of intentional understanding. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol. 33). Oxford: Elsevier.
- Woodward, L. A. (2009). Infants' Grasp of Others' Intentions. *Current Directions in Psychological Science*, 18(1), 53-57. <https://doi.org/10.1111/j.1467-8721.2009.01605.x>

**Notes**

Note 1. One could perhaps claim that this type of belief is no less integral to thinking in, say, quantum physics. The crucial point, here, is what I am about to say next, concerning the status of these beliefs even when not treated as precursors to Theoretical thinking.

Note 2. Other philosophers attributed the failure to find defining conditions for arthood to weaknesses in the suggested theories themselves (e.g., Bond, 1975, p. 177).

Note 3. This question was extensively discussed at the 2012 Balzan workshop on “Drama as experiment and Literature as an object of knowledge” (Murray Edwards College, Cambridge. January 2012) and the 2013 Balzan Colloquium on “Thinking with Literature” (CSMN, University of Oslo, September 2013).