

## Style Brings in Mental States<sup>1</sup>

### I

Much of what I write these days is an elaboration of Alan Palmer's argument that "novel reading is mind-reading." Here, too, I take up one aspect of that argument and consider it in light of my recent experience of studying fiction in a lab with fMRI equipment. (Well, not really—we are actually very far from approaching actual works of literature with brain imaging technology—but as close as I have come to doing so.)

As Palmer observes in his target essay for this volume, to claim that we understand the actions of fictional characters by uncovering "the mental network" behind them, is not to

flatten out the undeniable differences between novels, or to make impossible any worthwhile distinctions between them. To say that the reader can only follow the actions of the characters in Henry Fielding's *Tom Jones* (1749) by following the thought processes behind those actions is certainly not to say that it is the same sort of novel as James Joyce's *Ulysses* (1922). Of course the two are different.

And again, later in the essay,

[An] understanding of characters' thought processes is as necessary for *Tom Jones* as it is for *Ulysses*. I cannot find any way of retreating from the universality of my claim. Equally, I do not see any way in which this claim is a refusal to acknowledge the astonishing and endless variety of narrative. To say so would be like suggesting that I am trying to flatten out fictional variation by pointing out that *Ulysses* and Dan Brown's *The Da Vinci Code* use exactly the same 26 letters of the alphabet!

Palmer identifies here what constitutes both a problem and a "major cultural studies research project." If fiction is *all* about mind-reading, then the burden is on us (i.e., on cognitive literary critics) to explain why reading *Ulysses* feels so strikingly different from reading *Tom Jones* (not to mention *The Da Vinci Code*!). What goes into the construction of this difference? To what extent is it intrinsic to the text, and to what extent is it a reflection of the values of a particular historical period, or of an individual reader's perspective?

Palmer begins to address such questions by pointing out that, “Fielding gives us much less of the workings of characters’ minds than does Joyce, and so events are more central to the plot of the former’s novel and thoughts more central to the plot of the latter’s.”<sup>2</sup> He further observes that there seems to be significant variation in the way writers belonging to different historical periods construct “social minds” in the novel, a reflection, perhaps, “of the relationship between narrative technique and cultural conceptions of the self.” In what follows, I, too, argue that, far from “flattening out the fictional variation,” research in theory of mind may actually shed a surprising new light on how we construct such variation. Specifically, by becoming consciously aware of the “mental network” behind characters’ actions, we may gain a new appreciation of what constitutes an individual writing style.

## II

Style brings in mental states. That’s what I learned last summer, though my actual phrasing at the time reflected frustration rather than the joy of discovery: style *drags in* mental states.

As part of a research team, comprising literary scholars and cognitive neuroscientists, studying theory of mind with fMRI, I was in charge of putting together a series of narrative vignettes containing different levels of what we called “mental embedment.” To briefly illustrate our principle of counting levels of mental embedment, consider the following four examples: The sentence, “My last name begins with a Z, while Alan’s last name begins with a P,” contains no mental states, hence zero embedment. “I don’t want to read *The Da Vinci Code*” contains one mental state, that of *not wanting* to read the book, hence one embedment. “I used to think that I would hate *The Da Vinci Code*” contains two embedded mental states: *thinking* about *hating* the book. Finally, “I would have hoped that Alan could surmise that I wouldn’t want to read *The Da Vinci Code*,” contains three embedded mental states: my *assumption* about Alan’s *perception* of my *intentions*. (Note that the number of mental states doesn’t have to equal the number of people. The fifth-level mental embedment can be reached by having a single person reflect on her past and future states of mind.)

These are not actual sentences from the vignettes used in the study. According to the agreement we have with our cognitive scientist colleagues, we are not at liberty to quote any of the vignettes or to discuss the experiment itself before it is completed and the results, whether positive or negative, are in. I can, however, speak about the role of embedded mental states in fictional narratives—because it’s a topic that’s been central to my work for some time—and about the significance of the frustrating yet intriguing discovery—i.e., style *drags in* mental states—that I made in the process of editing vignettes.

First, the history of my discovery. To make the task of writing close to a hundred vignettes easier, we agreed to split it among several people. I hated writing mine. It's excruciatingly tedious to compose dozens of short narratives while strictly accounting for the number of mental states in each sentence of each narrative. For instance, "My last name begins with a Z, while Alan's last name begins with a P" might be a fine opening sentence for the "zero mental states" condition, but how do you develop this "story" from here if you are still not allowed to use any mental states? "In English, Alan's name would precede mine on any list, but here, in Russia, mine comes first." Seems all right, but isn't there a whiff of a mental state creeping in with, "here, in Russia, mine comes first"? The person who says it appears to be feeling *something*: perhaps, a satisfaction at being first in at least in some language, after a lifetime of bringing up the rear. So this sentence won't work. Anything that can be interpreted as introducing even a slight possibility of an extra mental state into a condition that may only have a certain number of mental states has to be excised.

So when I was done with my vignettes, they were all perfect images of the process that gave birth to them: boring. Here is a vignette similar to one of the vignettes from the actual study in the "zero mental states" condition:

I went grocery shopping two days ago. In the produce section, I got tomatoes, avocados, spinach, cilantro, green onion, and cauliflower. I also got fruit: apples, strawberries, grapes, and a watermelon. In the dairy section, I got milk, eggs, cheese, and yogurt. In the meat section, I got flounder and ground beef. I also bought olive oil, vinegar, dry beans, canned sardines, and paper towels.

Then I turned to the vignettes sent in by my colleagues. At first, I loved them. They were tacitly allusive and slyly imaginative; they read well (certainly much better than mine!); each had a sparkle. Once I started editing them, however, to check if there were any extra mental states, I realized that sparkles come at a price.

Not that there were any obvious *errors*, such as extra mental states expressed propositionally, as in "I thought," "she wanted," "he hated"—no, there was nothing as straightforward as that. Instead, various figures of speech introduced, now a touch of a thought, now a ghost of a desire, now an intimation of an attitude. Here is one such vignette, by Robert Barsky, who has generously agreed to let me quote it in this paper. Below I highlight in bold the elements of style that I see as creating mental states:

I drove the motorcycle **half-way across the country, alone**. I saw **21 states, 6 sunrises and 6 sunsets**, while **quietly sitting on a seat** that measures no more than **fourteen inches across**. The seat is a **magic carpet** that **whizzes through space** at 80 miles per hour, **supported by two wheels and an aluminum frame**. **I was connected to America**

**through that seat, and those little foot pegs that jut out from the lower part of the bike. Motorcycling is high-speed meditation.**

You may think that the problem with this vignette is that it has some first-level embedments—such as, “the speaker *feels* connected to America through his seat”—but it’s worse than that. The vignette actually has a *theme* running through it—the opposition between immobility and high speed, between the confines of the small seat and the little foot pegs and the vast expanse of the country—a theme that is developed tacitly throughout and then brought out explicitly in the concluding sentence: “motorcycling is high-speed meditation.”

So what happens to your theory of mind when you become aware of this theme as opposed to just taking in the factual information about the trip? As you realize that the vignette says something other than it seems to say at first glance, you start factoring in several minds: that of the speaker and that of the implied reader. For instance, does the speaker mean us to become aware of the opposition between the immobility and high speed even before the last sentence clinches it? And does he mean for us to notice that the vignette evokes the tradition of American road trip writing (which may entail re-experiencing what we felt while reading other works in that tradition)? And—an insight that I owe to Michael Holquist, the mastermind behind our interdisciplinary endeavor—what does the “eruption” of the “aphoristic general truth”<sup>3</sup> at the end of the vignette (i.e., “motorcycling is high-speed meditation”) do to our mind-reading? As with other truths “universally acknowledged,” do we instinctively respond with “says who?” and redouble our search for the mind behind this sentiment?<sup>4</sup>

In other words, if we are talking about underlying themes, evocations, and aphoristic truths, we are in at least the second level of embedment, for, to take in this vignette we have to become aware of the speaker’s anticipation of a certain response on the part of the reader. For instance, the reader might feel both jolted by the eruption at the end yet somehow prepared for it, and, by reckoning what it was that has prepared her for it, become increasingly conscious of the intentions of the speaker.

Of course, at the time, I didn’t think about themes and aphorisms. Busy as I was, I only noticed that the vignettes such as the one above evoked in me complex mental states even if these mental states were too tacit to describe quickly and in a propositional format.

At first I thought I could get rid of tacitly implied mental states while retaining the factual information contained in each vignette. But very soon I realized that a quick edit wouldn’t do: metaphors, aphorisms, and allusions were too deeply

integrated into the texts. I wrote to my fellow literary critics asking them to “dumb down” their vignettes, to make them bland, boring, unimaginative. We were in for a long process: Professors of literature have a difficult time *not* writing well, and elements of style drag in mental states.

### III

To give you a better idea why I consider this an exciting discovery, let me spell out the assumptions that underlie my thinking about theory of mind and fiction (and make me admire Palmer’s succinct formulation “novel reading is mind-reading”).

Theory of mind evolved to track mental states involved in real-life social interactions. On some level, however, our theory-of-mind adaptations do not distinguish between the mental states of real people and of fictional characters. Fiction, thus, feeds our theory of mind, giving us carefully crafted, emotionally and aesthetically compelling social contexts shot through with mind-reading opportunities. Hence the pleasure afforded by following minds in fictional narratives is to a significant degree a *social* pleasure. It’s an illusory but satisfying confirmation that we remain competent players in the social game that is our life.

Elsewhere I introduce a term *sociocognitive complexity* to discuss patterns of embedment of mental states within mental states, and I argue that a succession of scenes featuring third-level embedment—a mind within a mind within a mind—is the baseline for fiction. No fictional narrative can function on a lower level of sociocognitive complexity (though some experimental narratives try disguising mental states). Some authors/genres/works occasionally operate on the fourth level, and some reach even to the fifth and even sixth levels. In contrast, encyclopedia entries never rise to the third level, unless they deal with subjects that come with their own higher sociocognitive complexity (e.g., a Wikipedia entry featuring the plot synopsis of a novel or a movie).<sup>5</sup>

Until recently I focused my research on the two types of fictional mental states: those expressed propositionally and those inferred from observable body language. For example, Jane Austen frequently employs a combination of the two. She tells us what her characters are thinking/feeling (or, to adopt Palmer’s invaluable insight: what *intermental units* are thinking/feeling), and she lets us see characters’ bodies express their feelings, sometimes against their wills: “While [Elizabeth] spoke, an involuntary glance showed her Darcy, with a heightened complexion, earnestly looking at her, and his sister overcome with confusion, and unable to lift up her eyes” (204).

So, whereas I took for granted that there are other, numerous ways of introducing mental states into narratives, I never looked at them closely. That is, I intuitively

assumed that a paragraph-long list of inanimate objects encountered in a work of fiction can in principle be as powerful a conduit of sociocognitive complexity as a sentence of Austen describing Mr. Darcy's perception of Elizabeth's consciousness of his sister's discomfort, but I never thought through the implications of this assumption. Being confronted with vignettes which contained no propositional references to thoughts and feelings and no description of body language, yet still smuggled in extra mental states, made me aware that I was missing out on an important aspect of mind-reading in fiction: mental states brought in by what we broadly call elements of style, including figures of speech. (One exciting corollary to this new awareness was the realization that one can apply research in theory of mind to the study of poetry—something I've been wary of: poetry is often inimical to propositionally expressed mental states.)

Once more: the fMRI experiment that I am part of does not deal with excerpts from existing works of literature. The vignettes that I had to edit were relatively primitive synthetic constructs. And even there I found it impossible to get rid of extra mental states by simply dropping figures of speech. I discovered that a single metaphor can reorder the whole vignette, drawing other elements of style into its orbit to create a very particular tone. Revising would thus amount to writing an essentially new vignette.

Now think how this effect is exacerbated with actual works of fiction. After all, writers don't construct crudely obvious mental embeddings, such as "he *wants* them to stop *thinking* whatever they are *thinking* and *imagine* instead that they are in this place that he is *thinking* about." They may say instead, "Once more. Say you are in the country." Eight words, not a single direct reference to mental states, yet at least three embedded mental states. Paraphrase it, getting rid of Melville's style, and you may end up with zero mental states: "Ishmael is now talking about the country." This is, in effect, what study guides, such as SparkNotes do, as they dispense with the individual writing style of an author and hence downgrade the level of sociocognitive complexity of the original.<sup>6</sup>

Consider three more quotations:

The Reader may remember, that Mr. Allworthy gave Tom Jones a little Horse, as a kind of smart Money for the Punishment, which he imagined he had suffered innocently. (142)

He stayed in his walk to watch a typesetter neatly distributing type. Reads it backwards first. Quickly he does it. Must require some practice that. mangiD kcirtaP. (122)

Fache was in utter incomprehension of this woman's gall. Not only had she just barged in on Fache without permission, but she was now trying to convince him that Sauniere, in his final moments of life, had been inspired to leave a mathematical gag? (79)

I quote from the novels mentioned by Palmer—*Tom Jones*, *Ulysses*, and *The Da Vinci Code*—to make the three following observations. First, we can indeed only make sense of the characters' actions by "following the thought processes behind those actions" (Palmer). Second, each of these passages embeds three mental states, which, as I argue, constitutes the baseline level of sociocognitive complexity in fiction. Third, each passage builds sociocognitive complexity using different elements of style.

For instance, Fielding creates mental embeddings by factoring in mental states of the Reader and the Narrator: the Narrator *wants* the Reader to *understand* that Allworthy *believes* that Tom has suffered innocently. In contrast, Brown creates mental embeddings by focusing exclusively on the minds of his characters: Fache *can't understand* why Sophie *wants* him to *believe* that Sauniere spent the last moments of his life composing a mathematical gag. A Reader and a Narrator would be unthinkable in the sociocognitive ecology of *The Da Vinci Code* (and so would be—a brief glance at *Ulysses* here—free indirect discourse), not least because the author makes sure that there is no place for innuendoes and ambiguities in his representation of the characters' mental states. Hence in the passage above, Brown intensifies this effect by the strategic placement of such terms as "utter incomprehension," "gall," and "barge," which are used to create an impression that we have a fully measured and exhaustive account of Fache's strong feelings.

Which raises a question, in the words of Doug Whalen, the principal investigator on the cognitive-neuroscience side of our project, what happens when the reader "insists on being written into a text even when the text does not call any attention to the reader's existence."<sup>7</sup> We add mental states to the text by making it subject to our critical or classroom discussion; we modify the levels of embedment by modifying the context in which the text is read.

Viewed from this perspective, even the grocery-shopping vignette above can acquire unexpected mental states. Imagine this vignette as part of a novel. Its very inanity would be perceived as stylized and thus working toward particular narrative ends. We will be talking, for example, about the speaker's "flattening of affect" and wondering what caused this particular mental state in that character. In other words, what passes for the absence of mental states in the context of one genre (or, by extension, historical period) may acquire sociocognitive complexity when read within the context of a different genre (or a different historical period.)

What this all adds up to is that making strong "universal" claims about theory of mind and fiction hardly "flattens out variation"—instead, it forcefully focuses our attention on particularities of individual writing style and the context in which the text is read. If we know that works of fiction strive toward sociocognitive

complexity—that is, toward regularly embedding at least three mental states—we can ask how a specific text achieves it. What elements of style—used in an endless variety of idiosyncratic combinations—make a given sociocognitive complexity possible? And how can the same elements of style be read as adding up to a somewhat different configuration of sociocognitive complexity in a different cultural context?

### Notes

<sup>1</sup> I am grateful to the Teagle Foundation, to the members of Haskins-Teagle Collegium, and to the participants of the conference sponsored by the Foundation and held at the Haskins Laboratories on May 7, 2011: Doug Whalen, Michael Holquist, Robert Barsky, Evelyn Ender, Philip Rubin, Kenneth Pugh, Michael Warner, Blakey Vermeule, Alan Richardson, Gabrielle Starr, and Peter Steiner.

<sup>2</sup> For an important related analysis of mind-reading in Fielding, see Vermeule, “God Novels.”

<sup>3</sup> Michael Holquist, email communication, May 18, 2011.

<sup>4</sup> For a discussion of mind-monitoring involved in such “universally acknowledged” truths, see Zunshine, *Why We Read*, 62.

<sup>5</sup> See Zunshine, “What To Expect.”

<sup>6</sup> For a further discussion, see Zunshine, “What to Expect.”

<sup>7</sup> Doug Whalen, email communication, May 17, 2011.

### Works Cited

- Austen, Jane. *Pride and Prejudice*. Ed. James Kinsley; introduction by Fiona Stafford. New York: Oxford UP, 2004. Print.
- Fielding, Henry. *History of Tom Jones, a Foundling*. With an introduction and commentary by Martin C. Battestin, the text edited by Fredson Bowers. 2 vols. Middletown, Conn.: Wesleyan UP, 1975. Print.
- Joyce, James. *Ulysses*. New York: The Modern Library, 1961.
- Vermeule, Blakey. “God Novels.” In *The Work of Fiction: Cognition, Culture, and Complexity*. Eds. Alan Richardson and Ellen Spolsky. Aldershot: Ashgate, 2004. 147-166. Print.
- Zunshine, Lisa. “What To Expect When You Pick Up a Graphic Novel.” *SubStance: SubStance*. Special issue on graphic narratives. Ed. Jared Gardner and David Herman, # 124, 40.1 (2011): 114-134. Print.
- . *Why We Read Fiction: Theory of Mind and the Novel*. The Ohio State UP, 2006. Print.