Lisa Zunshine

Swift, Fielding, Johnson, Edgeworth: Cognitive Alternatives to Interiority

Draft—Please don't quote without permission

This essay argues that what we call the interiority of a fictional character is a cognitive-historical construct. That is, it emerges out of the interaction between our *theory of mind*—our evolved cognitive adaptation for explaining people's behavior in terms of their mental states, such as thoughts, desires, feelings, and intentions—and historically contingent ways of describing and interpreting behavior and mental states. At no point can the cognitive be separated from the historical, although various interpretive systems have considered and continue to consider them in isolation. Thus traditional literary analysis has been mostly unaware of the cognitive aspect of fictional interiority, while some overzealous evolutionary literary critics pointedly ignore the role of history in literary endeavor.[1] My goal here is to posit an interpretive model that is grounded in the workings of our cognitive adaptation for reading mental states into behavior and that is sensitive to particular cultural preoccupations of a given historical period. Specifically, I show how the eighteenth century's uneasy fascination with hypocrisy and madness informed the writers' intuitive reliance on their readers' cognitive predispositions as they worked to create an illusion of interiority of their socially troubling and personally troubled characters.

I

As an introductory illustration of the ongoing interplay between the cognitive and historical, consider a conversation that takes place in Laurence Sterne's *Tristram Shandy* (1759-67), when Walter Shandy, Uncle Toby, Yorick, and corporal Trim gather one evening around the fire in Shandy-hall. Yorick has just read out loud a long passage from Rabelais' *Gargantua*, in which Gymnast and captain Tripet turn somersaults on the croup of a horse, and now Walter prepares to share with the company his own manuscript, *Tristapaedia*:

> I have advanced nothing in *Tristapaedia*, but what is as clear as any one proposition in *Euclid*.—Reach me, *Trim*, that book from off the scritoir:—it has often been in my mind, continued my father, to have read it over both to you, Yorick, and to my brother Toby, and I think it a little unfriendly in myself, in not having done it long ago:—shall we have a short chapter or two now, and a chapter or two thereafter, as occasions serve; and so on, till we get through the whole? My uncle Toby and Yorick made obeisance which was proper; and the corporal, though he was not included in the compliment, laid his hand upon his breast, and made his bow at the same time.—The company smiled. *Trim*, quoth my father, has paid the full price for staying out the *entertainment*.—He did not seem to relish the play, replied Yorick.—'Twas a Tom-fool-battle, an' please your reverence, of captain *Tripet's* and that other officer, making so many summersets, as they advanced;—the *French* come on capering now and then in that way,—but not quite so much.

My uncle Toby never felt the consciousness of his existence with more complacency than what the corporal's, and his own reflections, made him so at that moment;—he lighted his pipe,—Yorick drew his chair closer to the table,—Trim snuff'd the candle,—my father stir'd up the fire,—took up the book,—cough'd twice, and begun. (321)

Why does Trim snuff the candle? We make sense of his behavior by exercising our theory of mind, that is, by attributing to Trim a certain intention: he *wants* to make the flame brighter so that Walter Shandy can see the pages of *Tristapaedia* better. In fact, we can read an even more complex mental state into Trim's action: He has just had a chance to voice his disapproval of the "battle" described in *Gargantua*, and he is flattered to be part of the discussion particularly as it seems to draw on his military expertise, even if Walter has not initially paid him the "compliment" of having wanted to read *Tristapaedia* to him as well as to Yorick and Toby. So by snuffing the candle Trim is not only making the pages more visible for Walter but also signaling his readiness to be included in this conversation and his satisfaction at having his opinion considered and respected by his "betters."

But wait! Aren't Trim's actions in fact incomprehensible? If he wants to make it easier for Walter to see the pages of *Tristapaedia* and also wants to signal that he belongs to this conversation, why would he snuff—that is, put out—the candle? As it turns out, the meaning of the word *snuff* has changed over the last two centuries. For Sterne's contemporaries, *snuffing the candle* meant "brightening its flame by freeing the candle from its excess wick, either by pinching or cutting off its snuff (the part of the wick

partially consumed in burning)."[2] This is to say that, prepared as we are to explain the characters' observable behavior in terms of underlying mental states, our explanations can be woefully wrong, if not downright impossible, unless we also know something about the historically contingent linguistic environment in which a given text was written and first read.

And the problem with words' changing their meanings as time goes by is, of course, only the tip of our interpretive iceberg. To get a fuller picture of the feelings animating the behavior of the protagonists in this passage, we have to know, for example, that the reason Yorick quotes a passage from *Gargantua*, in which Gymnast and captain Tripet do acrobatics in lieu of fighting, is not to comment on the effectiveness or futility of a particular way of waging the battle—which is how Trim takes it—but to satirize the practices of English "polemic divines" (319).[3] Armed with this literary-historical insight, we begin to appreciate the complexity of mental states underlying the nondescript actions of Toby as he lights his pipe, of Yorick as he draws his chair closer to the table, and of Walter Shandy as he stirs up the fire:

Toby is happy about having humored his own comrade (for it was Toby who, knowing that Trim would enjoy "the description of a battle," invited Trim to join them just as Yorick was about to read the passage from *Gargantua*). Yorick and Walter are sharing a joke about the similarities between the arguments of the "polemical divines" and the antics of Gymnast and captain Tripet—a joke of which Trim is not aware at all (and they know it), and which Toby appreciates only to a point because he is more happily preoccupied with Trim's satisfaction about voicing his opinion of that "Tom-fool-battle" and about being thus included in the conversation. Finally, Yorick must also

be pleased about humoring his host as Walter is about to read "a chapter or two" from his *Tristrapaedia*.

You notice that to make sense of this passage, we speak of Trim, Yorick, Walter, and Toby as endowed with interiority, that is, capable of a rich variety of mental states. This practice of treating fictional characters like real people has been long decried by literary critics, who see it as corrupting discussions both in classrooms and in non-academic reading environments, such as book clubs.[4] With the advent of research in theory of mind, I think it's time for us to admit that we've been fighting a "Tom-fool-battle" of our own (and have been cheating at it, too), reconcile ourselves to the fact that all of us, undergraduates, professors of literature, and book-club devotees, treat fictional characters as real people, and focus instead on historically specific strategies used by writers intuitively to exploit this "weakness" of their readers.

To clarify the role played by theory of mind (also known as *mind-reading*—clumsy terms both, but we are stuck with them for now) in our perception of fictional characters, consider this. After a certain age, people "cannot turn off their mind-reading skills even if they want to. All human actions are forevermore perceived to be the products of unobservable mental states, and every behavior, therefore, is subject to intense sociocognitive scrutiny."[5] The results of this "scrutiny" are far from perfect: we misread and misinterpret minds all the time, and we remain subject to people's conscious and subconscious manipulation of our perception of their mental states.[6] Still, flawed as our mind-reading is, it's the default way by which we construct and navigate our social environment. When theory of mind is impaired, as it is in varying degrees in the case of autism and schizophrenia, communication breaks down.[7]

So important is mind-reading for our species, that, at least on some level, we do not distinguish between attributing states of mind to real people and attributing them to fictional characters. Figuring out what an attractive stranger encountered by a fictional protagonist in a bookshop is thinking feels almost as important as figuring out what a real-life attractive stranger is thinking as she looks us in the eye and holds forth on how she enjoyed reading the book that we currently have in our hands. Thus the pleasure afforded by following various minds in fictional narratives is to a significant degree a *social* pleasure—an illusory but satisfying confirmation that we remain competent players in the social game that is our life. (Not surprisingly, patients with theory-of-mind deficits, such as those with autism-spectrum condition, exhibit a striking lack of interest in stories about people, both fictional and non-fictional.[8])

Talking about fictional characters as if they were real people is thus not a sign of the lack of readerly sophistication, but simply the reflection of the fact that the same cognitive adaptations that evolved to process information about mental states of real people also process information about mental states of fictional characters. Literary scholars are just as guilty of this offense as are lay readers. The main difference between "us" and "them" is that we can consciously adjust certain aspects of our mind-reading when we deal with a work of fiction, indeed may *enjoy* adjusting them, as a function of our personal mind-reading idiosyncrasies translated into the career of literary critic (i.e., a professional mind-interpreter).

That is, on the one hand, we attribute "real" mental states to characters with the same gusto as do our undergraduates. Think: when five minutes ago, you read the sentence, "Reach me, *Trim*, that book from off the scritoir," did you really pause to

remind yourself that you were dealing with a fictional construct, "Walter Shandy," who is portrayed so as to create an impression in the reader that he "wants" (well, actually "he" should be in quotation marks, too) another fictional construct, "corporal Trim," to "do" something for "him"? I don't think you did. I am sure that you simply skimmed over that sentence without even noticing that your theory of mind adaptations interpreted it as "Walter wants Trim to get him the book off the scritoir"—which is almost exactly the same kind of work that these adaptations would have performed had I, in your hearing, asked my daughter to reach me a book off the shelf.

On the other hand, more so than lay readers, we are apt at adjusting our mind-reading mode to introduce mental states of other "people" into our discussion of the novel: those of the novel's author, its actual and/or implied narrator, its variously historically situated readers, as well as living and dead literary critics and philosophers. For example, we may ask what was Sterne's own position on polemic divines? (That is, what did he *think* about them?) Or, what does Tristram, the novel's narrator, *mean* when he says that Yorick "read, or pretended to read" that long passage from Rabelais? What effect does this little interjection—"pretended to read"—have on contemporary readers, and what effect might it have had on Sterne's eighteenth-century readers? (That is, what do readers *think* about it?) What would Viktor Shklovsky say/think about it? What would Freud? What would Wittgenstein? What would Eve Kosofsky Sedgwick?

When we ask our students to consider such questions, it may seem that we encourage them to move away from treating fictional characters as real people and train them instead to see the novel as a historically situated artifact that uses a variety of narrative techniques to engage with ideological, aesthetic, and psychological agendas—

techniques that may come into sharper focus if we consider them via conceptual frameworks developed by various classical and modern thinkers. And train them we do. Make no mistake, however: The only way by which we can achieve this is by expanding the circle of entities whose minds we read *as if they were real people.*

This is what I meant, by the way, when I said earlier that, when we castigate lay readers for treating fictional characters as real people, we cheat (not intentionally, of course). Not only do we ourselves treat fictional characters as if they were capable of a broad variety of mental states (as real people are) to make sense of the story when we first read it; not only do we casually refer to these characters' and the author's mental states in our subsequent discussions with students (excepting, of course, those occasions on which we pointedly stop in our tracks and berate *them* for saying that Hamlet *wanted* that or that Jane Austen *intended* this); not only do we introduce more "people" into our conversation, such as Freud, Derrida, and Marx, and talk about them as if we know what they would think about the text; but we also cannot produce a single interpretation of the text—be it based in psychoanalytic, reader-response, feminist, or postcolonial theory— without constructing ever more elaborate and/or surprising attributions of thoughts, feelings, and intentions. Every single one of our exegeses involves attribution and interpretation of mental states, and none of our classroom conversations or scholarly publications would be practicable if we had to stop every minute and issue a disclaimer to the effect that we *seem* to be talking about real minds but we are really talking about fictional constructs.[9]

(Remember, too, that our much-touted contrast between the minds of real people and fictional characters is somewhat of a sham. We represent to ourselves the interiorities

of people around us using the same imperfect mind-reading adaptations that we use in representing interiorities of fictional characters. Indeed, I can say that there are some novel protagonists out there whose minds I seem to know better than I do the minds of some of my acquaintances. This might be a sign of the greater cognitive complexity of the latter, but it may also be a sign that I spend significantly less time thinking about their interiorities than I do about those of the former. And if that's the case, who is to tell me that the minds of the fictional protagonists are less real to me than the minds of my acquaintances, which I cobble together based on various external cues? Each real-life subjectivity is as flimsy a construction as a fictional one, if not flimsier.[10])

What all this adds up to is that the cognitive perspective on fictional interiorities commits us, *more than any other theoretical perspective*, to historicizing. As an evolved cognitive adaptation that has been with us for hundreds of thousands of years,[11] theory of mind does not change across the whole species just because a particular group of readers (say, eighteenth-century European middle-class devotees of sentimental novels, or twenty-first century North-American professors of literature and their students) happen to be in an environment that calls for a particularly elaborate attribution and interpretation of fictional minds. It follows then that a literary text from the tenth century BC (such as parts of the Hebrew Bible), a novel written between the second and third century A.D. (such as Heliodorus's *Aithiopika*), a Norse saga from the fourteenth century A.D., Sterne's *Tristram Shandy*, and Zadie Smith's *On Beauty*, all feed and stimulate the same enduring sociocognitive need to attribute and interpret mental states. But if they *all* do it, we have no way of explaining what feels like a very palpable difference in representation of fictional interiority originating in different time periods and cultures. That is, we have

no way of explaining it unless we focus on specific historical circumstances that encourage—or discourage!—the production and consumption of different modes of mind-reading in fictional representations.

In the rest of this essay, I discuss one strategy used by eighteenth-century writers to construct fictional interiority, a strategy that both reflected and critiqued the period's obsession with the mutability of the social self. The reason that I had to start with the long introductory discussion on theory of mind is that we can now talk about this particular strategy without making it bear more interpretive weight than it ought to bear. That is, instead of saying that the advance of this strategy heralded a broader cultural shift in the perception of subjectivity (a kind of claim routinely made by scholars of every literary period), I can advance a more modest and defensible argument. I can say that this strategy was just one of many ways in which eighteenth-century writers intuitively built on their readers' evolved cognitive eagerness to read mental states into behavior; that they did it by negotiating some of the period's immediate cultural preoccupations; and that, by developing a new way of representing fictional interiority they did not change their readers' mind-reading *adaptations* (which, again, remained constant for hundreds of thousands of years and will remain so for the foreseeable future), but they certainly changed their readers' mind-reading *environment*, as well as that of future readers and writers.

In other words, the grounding assumption of my argument is the refusal to overprivilege *or* to downplay either of the three views: 1) that fictional interiority resides in the text; 2) that it is imported into the text from a surrounding culture (a view that implies the reciprocal importation from the text to the culture); and 3) that it is something

that neurotypical readers (as opposed to people with autism-spectrum condition) cannot help but to read into the text as soon as they come across any description of behavior. The cognitive-historicist framework advocated by this essay presupposes a constant give-and-take among the text, the culture, and the mind, with interiority as a moving target, a shifting overlap of all three.

<div align="center">II</div>

Although as inveterate mind-readers we need only a mere whiff of a cue to see a mental state behind a behavior, it's been argued that one particularly socio-cognitively satisfying pattern of fictional mind-reading involves interaction among three mental states (no matter how many physical bodies these mental states may represent: one, two, or a crowd of a hundred and fifty).[12] Eighteenth-century writers explored a wide variety of social situations that allowed for such mental triangulations. In constructing such situations, however, they had to hierarchize the cognitive complexity of their characters, that is, they had to decide which characters should be able to represent more mental states than others.

For example, when the title protagonist of Eliza Haywood's *Fantomina* (1724), who impersonates different women to keep up the interest of her lover, Beauplaisir (who doesn't know that all of his mistresses are actually the same person), receives two letters from Beauplaisir—one addressed to "Widow Bloomer," his fresher conquest, whom he is eager to see immediately, another to his by now "stale" passion, "Fantomina," whom he

needs to put off for the sake of the Widow and to whom he thus lies about being busy—the mind-reading dynamics of the scene put her at the apex of its mental triangulation:

> Traytor! (cry'd she) as soon as she had read [both of Beauplaisir's letters], 'tis thus our silly, fond, believing Sex are serv'd when they put Faith in Man: So had I been deceiv'd and cheated, had I like the rest believ'd, and sat down mourning in Absence, and vainly waiting recover'd Tendernesses. . . . But I have outwitted even the most Subtle of the deceiving Kind, and while he thinks to fool me, is himself the only beguiled Person. (59)

Observe that Beauplaisir imagines two mental states—his own and the presumably fooled Fantomina—whereas Fantomina imagines three: her own, Beauplaisir's as he imagines the fooled Fantomina, and the presumably fooled Fantomina's (who is thus likened to "fond" and "silly" women who trust men). To put it differently, she is capable of representing the feelings of Fantomina as represented by Beauplaisir, thus engaging in what cognitive scientists call a third-level cognitive *embedment*. If we correlate fictional interiority with the ability to embed multiple mental states we begin to see the workings of the narrative mechanism that endows some characters with "more" interiority than others. At least for now, Fantomina is on top of the cognitive "food chain," and as such she is the most interesting character in the story.

I have argued elsewhere that the writer's intuitive decision about which of her characters will carry on complex mind-reading embedments and which will have to settle

for simpler ones can be informed by considerations of social class, of gender or race, or

of any other parameter reflecting current ideological investments of the society.[13] Here I

want to consider yet different aspect of that decision. The characters not capable of

representing anybody's mental states except their own, such as Lady Bertram from Jane

Austen's *Mansfield Park* (1814),[14] may come across as stupid or flat, their interiority at

the lowest ebb; but the characters capable of representing multiply embedded mental

states may come across as peculiarly misguided, betrayed as it were by their cognitive

complexity into ethically questionable or socially debilitating behavior.[15] It seems that

readers' perception of interiority imposes a cost on a fictional character, though the exact

form of this cost is figured on a case-by-case basis, reflecting, among other things,

specific cultural preoccupations of the moment. In what follows, I consider two models

of eighteenth-century fictional interiority—one constructed specifically via the

characters' capacity for embedding multiple mental states, and another constructed via

their capacity for considering perspectives differing from their own—and discuss the

costs associated with each model.


*Hypocrisy*


Eighteenth-century anxiety about the mutability and endless performativity of the

self, and the resulting uneasy view of hypocrisy as a social glue, are well-established

topics in recent scholarship, explored in such studies as Blakey Vermeule's *The Party of

Humanity: Writing Moral Psychology in Eighteenth-Century Britain* (2000), Patricia Ann

Meyer Spacks's *Privacy: Concealing the Eighteenth-Century Self* (2003), and Jenny

Davidson's *Hypocrisy and the Politics of Politeness* (2004). From the cognitive

perspective espoused by this essay, hypocrisy is a quintessential mind-reading

phenomenon: an endeavor to subtly convince another person that she knows your

thoughts and feelings and that these thoughts and feelings are congenial to her own. It

presupposes a third-level mental embedment ("I know that you like X, so I *want* you to

*think* that I *like* X, too) and as such presents a particularly inviting subject for fictional

and non-fictional explorations.

Looking at how the eighteenth century treated its hypocrites, one gets an

impression that non-fictional discourses were more forgiving than fictional ones. Thus

*The English Theophrastus* (1702) refuses to pass judgment on hypocrisy, admitting

instead with world-weary resignation that not only is sincerity a particularly insidious

performance of openness ("Half-witted People can never be sincere . . . Sincerity . . . is

only a more cunning and shrewd sort of Dissimulation, to insinuate our selves into the

Confidence of other People"), but also that every social interaction involves mutual

manipulation ("All Men [strive to] appear what they have a mind to be taken for: So that

we may say, That the World is made up of nothing but formal Countenance and Shews"

[205]).

The Earl of Chesterfield's *Letters to His Son* (1774) is a heartfelt glorification of

social hypocrisy. Its author frankly delights in his capacity for adapting to the minds of

others and fervently wants his son to learn this social grace. Thus describing his younger

self: "My passion for pleasing was so strong (and I am very glad it was so), that I own to

you fairly, I wished to make every woman I saw in love with me, and every man I met

admire me" (1: 31). On strategies of making people like you: Be attentive to small things

that they value, for "attention to such trifles flatters self-love much more than greater things, as it makes people think themselves almost the only objects of your thoughts and care" (1:29). On pleasing people whose opinion you, in fact, despise:

As women are a considerable, or at least a pretty numerous part of company; and as their suffrages go a great way toward establishing a man's character in the fashionable part of the world (which is of great importance to the fortune and figure he proposes to make in it), it is necessary to please them.

However:

The man of sense only trifles with them, plays with them, humors and flatters them, as he does with a sprightly forward child; but he neither consults them about, nor trusts them with serious matters; though he often makes them believe that he does both; which is the thing in the world that they are proud of . . . [for] being justly distrustful that men in general look upon them in a trifling light, they almost adore that man who talks more seriously to them, and who seems to consult and trust them (2:107).

It's tempting to see the period's fictional narratives as offering a wishful corrective to successful careers of real-life hypocrites. (For the Earl of Chesterfield was a success, or, at least, considered himself such). Think of one assiduous third-level

embedder, Blifil, from Fielding's *Tom Jones* (1749), who endeavors to convince every person instrumental to his social advancement that he shares his worldview. Already at sixteen, he has "address enough . . . to recommend himself at one and the same time" to both his mentors, Thwackum and Square, in spite of their opposing views: "With one he was all religion, with the other he was all virtue." Moreover, by praising Thwackum and Square to Allworthy behind their backs, Blifil not only ensures that they would think that he sincerely values their instructions (for he knows that Allworthy repeats "all such compliments to the persons for whose use they were meant"), but that his uncle, too, would find such panegyrics "extremely grateful," as they so loudly resound "the praise of that singular plan of education" which Allworthy "laid down" for the boys (116).

A real-life Blifil would have been very likely to succeed in a world in which sincerity is "only a more cunning and shrewd sort of Dissimulation, to insinuate our selves into the Confidence of other People." Fielding, of course, doesn't let him get away with anything, finally unmasking him in front of Allworthy. True, Blifil doesn't end up completely destroyed—in fact, when we last see him he is clawing his way back to the top by using his old strategy of pleasing people who matter, such as turning Methodist "in hopes of marrying a very rich widow of that sect, whose estate lies in that part of the kingdom" for which he wants to "purchase a seat in the next Parliament" (869). Still, within the moral economy of the novel he remains an irredeemable villain.

Similarly, Thomas Holcroft's *Hugh Trevor* (1797) contains a panoply of third-level embedders none of whom suffer a complete worldly fiasco, but each of whom is intended to strike the reader as repugnant in his own way: from the reverend Enoch Ellis, whose life is one "necessary endeavor to please" (102); and Earl of Idford, who wants to

flatter a brilliant thinker and public speaker Mr. *** and thus proclaims—*when that gentleman is present*—that he "would rather be Mr. *** than a prince of blood" (133); to Lord Bishop, who, together with his friend, the Dean, sprinkle their dinner conversation with crude obscenities while "hypocritically [avoiding] words which the ear could not endure" (148), so as not to alienate completely the other people present. Hugh himself has to learn not to be "too eager in [his] thirst of approbation" (118), a thirst which leads him, among other things, to such repulsive displays of hypocrisy as complimenting the Bishop on his cathedral sermon (which Hugh himself wrote, but "his lordship" shamelessly passed for his own and, worse yet, completely mangled by his "spiritless" delivery, "thick" voice, and "ridiculous" cadences [138]).

It seems, in other words, that eighteenth-century fictional characters who practice third-level cognitive embedment for the purposes of social elevation get condemned as hypocrites and judged harshly, as opposed to characters who demonstrate an acute awareness of other people's mental states—which also frequently calls for a third-level mental attribution—but have no worldly ambitions and thus refuse to manipulate others by pretending that they share their feelings (e.g., Tom Jones himself, as well as Turl and Mr. Evelyn from *Hugh Trevor*). The novels' distaste for the mind-reading pattern (i.e., Theophrastian "sincerity" and Chesterfieldian "graces") that in real life could indeed be rewarded by social advancement is quite striking. Cognitive construction of fictional interiority becomes an ideological endeavor, enabling what John Richetti sees as the "deeply critical" view of "more traditionalist authors such as Fielding" of the "emerging modern self and the new kind of society that encourages it" (8).

*Madness*

Madness attracted and repelled eighteenth-century readers and writers as strongly as did hypocrisy, and as such it has received sustained attention from modern scholars, from Michel Foucault (*Madness and Civilization*, 1965) and Max Byrd (*Visits to Bedlam Madness and Literature in the Eighteenth Century*, 1974) to Clement Hawes (*Mania and Literary Style: The Rhetoric of Enthusiasm from the Ranters to Christopher Smart*, 1996) and Allan Ingram and Michelle Faubert (*Cultural Constructions of Madness in Eighteenth Century Writing*, 2005). To consider representations of madness from the cognitive perspective taken by this essay, we may ask to what extent was madness correlated with particular constructions of fictional interiority? To put this differently, is there a cognitive method to eighteenth-century madness, that is, will we understand better at least some forms of it if we adapt the vocabulary of mind-reading and embedded mental states?

One distinguishing feature of many eighteenth-century mad (or just eccentric in ways that strike observers as mad) fictional characters is their misplaced mind-attribution. That is, they attribute mental states either to entities that don't have any (such as natural phenomena) or to entities whose mental states should matter little or not at all in one's everyday life (such as other fictional characters or imaginary creatures). What is interesting about this "mad" behavior is that it builds on an essential aspect of our mind-reading endowment, which, under normal conditions, is crucial for our social functioning:

namely, the capacity for being aware of other people's perspectives, particularly when they differ from our own.

In fact, the deficits of this capacity manifest themselves early enough in the development to be used as the ground for diagnosing autism. The test administered to four-year olds (the age of four being an important threshold in the maturation of theory of mind) involves putting a child in a situation in which she has to demonstrate an understanding that other people may entertain false beliefs, that is, that they may believe that something that she *knows* to be the case is not so. Neurotypical children pass this "false belief" test, while children with autism-spectrum condition generally don't.

Hence one strategy for creating the illusion of interiority of fictional characters is to have them be aware of differing perspectives of the same subject. Maria Edgeworth uses this strategy in the opening paragraph of her *Castle Rackrent* (1800). The novel's narrator, Thady Quirk, tells us how different people in and out of the "family" view *him*: he is "honest Thady" to some, "old Thady" to others, "poor Thady" to yet others, Thady-to-be-disregarded to his own son, and "true and loyal" Thady in his own eyes:

> Having out of friendship for the family, upon whose estate, praised be
> Heaven! I and mine have lived rent free time out of mind, voluntarily
> undertaken to publish the Memoirs of the Rackrent Family, I think it is my
> duty to say a few words, in the first place, concerning myself.—My real
> name is Thady Quirk, though in the family I have always been known by
> no other than '*honest Thady*'—afterwards, in the time of Sir Murtagh,
> deceased, I remember to hear them calling me '*old Thady*;' and now I'm

come to '*poor Thady*'—for I wear a long great coat winter and summer . . . to look at me, you would hardly think '*poor Thady*' was the father of attorney Quirk; he is a high gentleman, and never minds what poor Thady says, and having better than 1500 a-year, landed estate, looks down upon honest Thady, but I wash my hands of his doings, and as I have lived so will I die, true and loyal to the family. (7-8)

Note that, defined in terms of the character's ability to consider different perspectives, fictional interiority does not necessarily translate into intellectual sophistication or a heightened capacity for empathy. That is, such a character may still come across as naïve (as Thady does; though some critics also read his naïveté as a clever mask adapted by a manipulative and "self-serving man"[16]) or as obnoxiously self-centered. Thus Lady Clementina from Elizabeth Inchbald's *Nature and Art* (1796) is acutely aware of other people's perspectives, yet all she can think of is what all these different people think about *her*:

If she complained she was ill, it was with the certainty that her languor would be admired; if she boasted she was well, it was that the spectators could admire her glowing health; if she laughed, it was because she thought it made her look pretty; if she cried, it was because she thought it made her look prettier still.—If she scolded her servants, it was from vanity, to show her knowledge superior to theirs; and she was kind to them for the same motive, that her benevolence might excited their

admiration.—Forward and impertinent in the company of her equals from

the vanity of supposing herself above them, she was bashful even to

shamefacedness in the presence of her superiors, because her vanity told

her she engrossed all their observation (50).

In other words, a character's relative cognitive complexity (that is, her ability to

embed mental states of numerous other people), which strongly contributes to our

perception of her interiority, can be disconnected from her intelligence and, in fact, from

any true awareness of other people's states of mind. A character can be an assiduous

mind-reader, yet a hopeless solipsist; "neurotypical" as far as her theory of mind goes, yet

lacking in empathetic imagination.

She can also be mad. The difference between such characters as Lady Clementina,

on one hand, and Jonathan Swift's Gulliver (*Gulliver's Travels*, 1726), Charlotte

Lennox's Arabella (*Female Quixote*, 1752), and Samuel Johnson's Astronomer

(*Rasselas*, 1759) on another, is that solipsistic as the former can be, they seek their

reflection in the minds of socially acceptable/sanctioned others, whereas the latter look

for it in the minds that are not "really" there. It's as if their theory of mind was working

in overdrive, not discriminating between those whose opinion about them should matter

and those who simply cannot have any opinion.

Take Gulliver. It's not just that he has what Michael Seidel calls "an image

problem," not knowing "whether he looks good small or big"; "whether features of his

face look better symmetrically disposed or awry"; "whether his pathetic body would

serve him better shaped like a horse's"; "whether his sounds and smells are as offensive

to everyone else as they are to him" (75)—after all, Lady Clementina has a kind of a "image problem" too—the trouble is that he turns for answers to these questions to the creatures who (according to one interpretive tradition[17]) might be mere figments of his imagination: tiny Lilliputians, giants from Brobdingnag, inhabitants of a flying island, and talking horses. In any case, it's his reflection in their minds that matters to him—to the exclusion of people whose opinion *ought* to matter. As he puts it in the letter to his Cousin Sympson: "The united Praise of the whole [human] Race would be of less Consequence to me, than the neighing of those two degenerate Houyhnnms I keep in my Stable; because, from these, degenerate as they are, I still improve in some Virtues, without any Mixture of Vice" (9-10).[18]

Just so Arabella from *Female Quixote* cares deeply about the opinion of phantoms but not about that of her actual acquaintances. In stark contrast to her cousin, Mr. Glanville, who writhes in embarrassment when he imagines what the people whom they meet, such as Sir George, must think of Arabella's odd behavior, Arabella is concerned with how she would look in the eyes of the characters from her beloved Romances; in the eyes of the historians who would come to write the "History" of her adventures; and in the eyes of the future readers of that "History." Thus, upon receiving a letter from Sir George, which, she suspects, contains a confession of love, she debates with herself whether she should read it and thus have her suspicion confirmed, which must necessarily result in her banishing Sir George from her presence forever. No Romance heroine that she can think of would read such a letter, hence "*fearful of transgressing the Laws of Romance*, by indulging a Curiosity *not justifiable by Example*, [Arabella resolves] to return this Letter unopened." Further deliberation on the matter even makes her angry,

and she speaks to the letter thus: "Presumptuous Paper! . . . Bold Repository of thy Master's daring Thoughts! *Shall I not be blamed by all, who, hereafter, will hear or read my History*, if, contrary to the Apprehensions I have, that thou containest a Confession that will displease me, I open thy Seal and become accessory to thy Writer's Guilt, by deigning to make myself acquainted with it?" (198; emphasis added)

Finally, the Astronomer from *Rasselas*, convinced that the sun listens "to [his] dictates," passing "from tropic to tropic by [his] directions," that the clouds, "at [his] call," pour their waters, and that the Nile overflows "at [his] command," is not merely attributing mental states to natural events (a phenomenon that cognitive psychologist Jesse M. Bering would see as building on our "existential theory of mind"). He also values the opinion of the nebulous entities who bestowed "this distinction" on him more than that of ordinary people, whose ridicule and incredulity he is prepared to disregard. Thus when he discovers in Imlac, who listens to his story, "some tokens of amazement and doubt," he observes that, "not to be easily credited will neither surprise nor offend [him]; for [he is], probably, the first of human beings to whom this trust has been imparted" (466).

These three cases of irrational mind-attribution are relatively clear cut: as long as Gulliver, Arabella, and the Astronomer privilege the good opinion of phantoms over that of the "real" people, they could be considered mad. (Arabella and the Astronomer eventually "get better" by learning to seek their reflections in the "right" minds.) Other fictional characters who construct their self-image by reading minds of others walk a thinner line. Members of their social circle may pronounce them insane when they chose to be reflected by the people whose thoughts about them should not matter—because of

their class, race, gender, or a particular ideological concern of the moment. After all, a Chesterfieldian "man of sense" (as opposed to what: a man who is silly? irrational? insane?) must not really care about what women think of his affairs—he must only make them think that he cares.

Portraying a fictional character as aware of her reflection in the minds of others is one strategy for making us perceive her as having interiority—a *cognitive* strategy. Deciding *whose* minds she should or should not be reflected in is an intensely ideological, culture-specific choice, made first by the writer but then—again and anew— by the reader situated in a particular historical moment. Fictional interiority is thus a social phenomenon, but as such it is always a cognitive-historical construct. Hence to Deidre Lynch's argument that "individuated psychological meanings . . . do not come naturally" to readers of any historical period, that they are instead "social productions, objects of contest as well as of collaboration" that "become intelligible in historically specific, institutionally mediated ways" (9), I would add that they are always cognitive productions as well.

---

Notes:

I am grateful to Clement Hawes for his insightful suggestions regarding the earlier draft of this essay and to Judy Prats for her valuable editing help.

[1] See Richardson, 12-13, and Spolsky.

[2] Melvyn New, "Notes," in Sterne, *Tristram Shandy*, 622, note 3.

[3] Ibid., note 6.

[4] See Long, 156, and Woloch, 16-17.

[5] Bering, 12.

[6] See Zunshine, "Theory of Mind."

[7] See Baron-Cohen; Frith.

[8] See Barnes; Barnes et al.

[9] For a related discussion, see Zunshine, *Why We Read Fiction*, 99-100.

[10] For a fuller discussion, see Palmer.

[11] See Baron-Cohen, 13.

[12] See Vermeule, "Machiavellian Narratives" and Zunshine, "1700-1775."

[13] See Zunshine, "1700-1775."

[14] See Zunshine, "Mind Plus."

[15] Compare to Vermeule's observation in "Machiavellian Narratives," that the ability to reflect other people's mental states does not automatically translate into superior ethics: crafty villains can be "masterminds" carrying on triple or even quadruple mental embedments.

[16] Bauermeister, 39.

[17] For a strong view of Gulliver's madness, see Seidel's "Introduction" (particularly "The Madness of Gulliver," xiii-xv); for an opposing view, see Rawson, 191.

[18] See Vermeule's "Satirical Mind-Blindness" for a compelling yet very different application of theory of mind to Swiftian satire.

Works Cited:

Barnes, Jennifer. "Fiction and Empathy: Narrative Cognition in Autism." Talk in the

series "Current Work in Developmental Psychology." Department of Psychology,

Yale University. January 14, 2009.

Barnes, Jennifer, R. Li, Simon Baron-Cohen, and Paul Bloom. "Reading Preferences,

Empathy, and Autism." In preparation.

Baron-Cohen, Simon. *Mindblindness: An Essay on Autism and Theory of Mind*.

Cambridge: The MIT Press, 1995.

Bering, Jesse M. "The Existential Theory of Mind." *Review of General Psychology* 6.1

(2002): 3-24.

Bauermeister, Erica, Jesse Larsen, and Holly Smith. *500 Great Books by Women: A

Reader's Guide*. Penguin, 1994.

[Boyer Abel]. *The English Theophrastus: or, the Manners of the Age. Being the Modern

Characters of the Court, the Town, and the City*. London: Printed for W. Turner,

at Lincolns-Inn Back Gate; R. Basset in Fleetstreet; and J. Chantry, without

Temple Bar, 1702.

Edgeworth, Maria. *Castle Rackrent*. Ed. George Watson. London: Oxford University

Press, 1964.

Fielding, Henry. *Tom Jones*. Ed. John Bender and Simon Stern. New York: Oxford

University Press, 1996.

Frith, *The Cognitive Neuropsychology of Schizophrenia*. Hove, East Sussex, UK:

Lawrence Erlbaum Associates, 1992.

Haywood, Eliza Fowler. *Fantomina and Other Works*. Ed. Alexander Pettit, Margaret

Case Croskery, and Anna C. Patchias. Broadview Press, 2004.

Holcroft, Thomas. *Hugh Trevor*. Ed. Seamus Deans. Oxford: Oxford University Press,

1978.

Johnson, Samuel. "The History of Rasselas, Prince of Abissinia." *Johnson Prose and

Poetry*. Ed. Mona Wilson. Cambridge, MA: Harvard University Press, 1967. 389-

482.

Lennox, Charlotte. *The Female Quixote*. Ed. Amanda Gilroy and Wil Verhoven. London:

Penguin Books, 2006.

Lynch, Deidre. *The Economy of Character: Novels, Market Culture, and the Business of Inner Meaning*. Chicago: University of Chicago Press, 1998.

Palmer, Alan. *Fictional Minds*. Lincoln: University of Nebraska Press, 2004.

Rawson, Claude Julien. *Swift*. London: Sphere, 1971.

Richardson, Alan. "Studies in Literature and Cognition: A Field Map." *The Work of Fiction: Cognition, Culture, and Complexity*. Ed. Alan Richardson and Ellen Spolsky. Aldershot: Ashgate, 2004. 1-30.

Richetti, John. "Introduction." *The Cambridge Companion to the Eighteenth-Century Novel*. Ed. John Richetti. Cambridge University Press, 1996. 1-10.

Seidel, Michael. "*Gulliver's Travels* and the Contracts of Fiction." *The Cambridge Companion to the Eighteenth-Century Novel*. Ed. John Richetti. Cambridge University Press, 1996. 72-89.

-----. "Introduction." Jonathan Swift, *Gulliver's Travels*. New York: Barnes and Noble, 2004. xiii-xxxi.

Spolsky, Ellen. "Frozen in Time?" (Review of Jonathan Gottschall and David Sloan

    Wilson, eds., *The Literary Animal: Evolution and the Nature of Narrative*)

    *Poetics Today* 28.4 (Winter 2007): 807-816.


Swift, Jonathan. *Gulliver's Travels*. Edited with an Introduction by Claude Rawson and

    notes by Ian Higgins. New York: Oxford University Press, 2005.


Vermeule, Blakey. "Satirical Mind Blindness." *Classical and Modern Literature* 22.2

    (2002): 85-101.


-----. "Machiavellian Narratives." *Introduction to Cognitive Cultural Studies*. Ed. Lisa

    Zunshine. Baltimore: Johns Hopkins University Press, forthcoming.


Woloch, Alex. *The One vs. The Many: Minor Characters and the Space of the*

    *Protagonist in the Novel*. Princeton University Press, 2003.


Zunshine, Lisa. *Why We Read Fiction: Theory of Mind and the Novel*. Columbus: Ohio

    State University Press, 2006.


-----. "Theory of Mind and Fictions of Embodied Transparency." *Narrative* 16.1 (2008):

    65-92.

-----. "Mind Plus: Sociocognitive Pleasures of *Northanger Abbey* and *Mansfield Park.*"

*Studies in Literary Imagination*, forthcoming.


-----. "1700-1775: Theory of Mind, Social Hierarchy, and the Emergence of Narrative

Subjectivity." *The Emergence of Mind: Representations of Consciousness in

Narrative Discourse in English, 700 – the Present*. Ed. David Herman. Lincoln:

University of Nebraska Press, forthcoming.