Abstract and Keywords

This chapter suggests that nested mental states constitute a fundamental unit of meaning in fiction. It uses the concept of sociocognitive complexity (mental state nested within mental state, and so forth) to offer new ways of reading Cao Xueqin’s classic novel *The Story of the Stone* (Honglou meng 紅樓夢). The chapter discusses the contrast between explicitly described mental states of characters and implied nested mental states of characters, narrators, implied readers, and authors, by bringing together research on theory of mind (particularly, in developmental psychology: Joan Peskin and Janet Wilde Astington); narratology (Robert Scholes, James Phelan, and Robert Kellog); and scholarship on *Hong Lou Meng* (Dore J. Levy, Andrew H. Plaks, David L. Rolston, Haun Saussy, Andrew Schonebaum, Angelina C. Yee, Anthony C. Yu, Kam-Ming Wong, and Xueping Zhong).

Keywords: sociocognitive complexity, 紅樓夢, Hong-Lou Meng, The Story of the Stone, Cao Xueqin, nested mental state, implied mental state, theory of mind, developmental psychology, fiction

AN old tree next to my house needs to be cut down, yet the contractor keeps postponing, and I am worried that yet another dead branch will fall on my neighbors’ car. I don’t know how they explain to themselves that I haven’t yet taken care of it. I want them to know that I am thinking about this issue. I shall email them.

I want them to know what I think—a mental state within a mental state within a mental state—three nested mental states. As I survey my day, more examples of such nestings from different occasions come to mind. She thought that I meant the opposite of what I actually meant. He didn’t want me to know what he was really thinking. I don’t want her to realize that I am trying out this new communication strategy that I just learned from a book. When he’s older, do you think he’ll forget how he felt when he was four?

It’s difficult to say how much of our daily functioning involves nesting mental states within each other in this recursive fashion (particularly since we don’t stop and think about it consciously the way I just did.)¹ It seems to me that we do it
often, though not constantly. Involved social situations call for at least some triply nested thoughts and feelings. Or, perhaps, involved social situations are *created* by our ability to entertain such nestings.

Emotions play a crucial role in nesting. To use just one example, it is possible that we nest mental states more actively in situations conducive to social anxiety. In fact, social anxiety may be viewed as an archetypal triple nesting: our awareness of our uncertainty about other people’s feelings in response to their perception of our actions.

That cognition should thus be tied with emotion is not surprising given the fraught nature of theory of mind—our evolved cognitive adaptation for attributing mental states to ourselves and others. To stay just with social anxiety, theory of mind comes all but preloaded with it because it makes us interpret all observable behavior as caused by underlying thoughts, feelings, and intentions, yet our interpretations are often wrong, and we know that they can be wrong.

Social failures, large or small, are thus always just around the corner. We may nest aggressively to forestall them, but it’s also possible that by doing so we give shape to our social anxiety and intensify it. I am unhappy that the contractor is not here yet. Oh, but what I am actually unhappy about is that I am afraid that my neighbors will think that I don’t care about their (very real) concerns. And I am sure that they don’t want to remind me because they don’t want me to think that they don’t trust me to take care of it... Am I being socially astute or am I overthinking it?

Fiction is where it gets really interesting. Nested mental states suffused with strong emotions are everywhere in fiction. Yet writers can construct them by referring to other mental states or without mentioning mental states at all. Hence Cao Xueqin’s *The Story of the Stone* (*Shitou ji* 石頭記) also known as *Dream of the Red Chamber* (*Honglou meng* 紅樓夢), a beloved Chinese novel written around 1760 and first published in 1792:

> And now suddenly this Xue Bao-chai had appeared on the scene—a young lady who, though very little older than Dai-yu, possessed a grown-up beauty and aplomb in which all agreed Dai-yu was her inferior.

What’s going on in this sentence? Here is one way to spell out the mental states that we infer as we make sense of it: the narrator *wants* his readers to *realize* that Dai-yu *feels distressed* because she *is certain* that everyone around her *considers* her inferior to Bao-chai. That’s at least four recursively nested mental states (and some serious social anxiety), but to articulat them, we have to take in subtle cues, such as the unhappy tone with which Dai-yu refers to her cousin (“this Xue Bao-chai”：一個薛寶釵) and our previous awareness of Dai-yu’s near-paranoid self-consciousness.

If we look for *explicit* references to mental states this sentence contains, we notice the word rendered by the translator as “agreed” (謂). This word may describe an attitude of some people around Dai-yu, but the meaning of the passage does not reside with it. Instead, as we’ve seen, that meaning is expressed through nested mental states implied but not stated by the text.

Literary critics have long known that fiction can represent mental states without referring to them. To quote Haun Saussy, “That thoughts can be represented even if unspoken is a commonplace of universal literature.” Here is what’s new about the cognitive approach that I propose here. I focus on *nested* mental states as units of meaning in fiction,
including situations when these nested mental states are implied rather than described.

Viewed from this perspective, fiction emerges as both continuous with our everyday cognition and distinct from it. Building on theory of mind and mimicking patterns of our everyday social functioning, fiction nests mental states within mental states. Yet fictional nestings are not merely crude transcriptions—along the lines of, “I think that he thinks that I think”—of the myriad subtle cues (particularly those conveyed by body language) that make up our social life. Instead, fiction creates complex mental states by stylistic means unique to specific genres and authors. Fictional nestings of thoughts and feelings have their own history—a literary history—not reducible to social cognition (however complex that is).

To talk about nested mental states in fiction, I introduced, elsewhere, the term sociocognitive complexity. I have argued that prose fiction, drama, narrative poetry, as well as memoirs concerned with imagination and consciousness (such as Nabokov’s *Speak, Memory*) routinely operate on at least the third level of sociocognitive complexity: a mental state within a mental state within yet another mental state. In this, they differ from our daily social interactions, which feature third-level nestings only sporadically, as well as from other ostensibly nonfictional discourses, such as newspaper articles (which operate comfortably on the second level) or science textbooks (which can get by without nesting any mental states).

Once you start reading a work of fiction, you encounter third-level nestings very soon and after a while are immersed in them. Different authors achieve this by different stylistic means, focusing primarily on mental states of either characters or of narrators, implied authors, and implied readers. For example, the sociocognitive complexity of Zamyatin’s *We* or McCarthy’s *Blood Meridian* depends on the nested mental states of their implied authors and readers to a much greater degree than does the sociocognitive complexity of J. K. Rowling’s “Harry Potter” series, which mainly nests the mental states of its characters.

Some writers operate on the fourth level of sociocognitive complexity, and some reach even to the fifth and sixth, though there is no clear correlation between those higher levels of sociocognitive complexity and the aesthetic value of the work.

Cognitive literary analysis thus brings into the open an important aspect of our engagement with fiction: we make sense of complex social behavior, inevitably present in fiction, in terms of nested mental states. But whose and where are those mental states? Are they objective, immutable, permanently ensconced in the text, ready to be discovered by whoever opens the book?

Once more, a parallel with real-life social interactions is illuminating here. In real life, understanding behavior as caused by nested (and thus complex) thoughts, desires, and intentions, always involves interpretation of those mental states. There may be no such thing as an objective, correct interpretation, unless the social context is drastically simplified. Similarly, in fiction, understanding what we read in terms of our own, the characters,’ and the (implied) author’s thoughts and feelings, inevitably involves constructing nestings that are subjective and context-sensitive.

Take *The Story of the Stone*. If it’s true that Red Inkstone and Odd Tablet, the influential early commentators who read *Stone* in manuscript, were indeed personally familiar with people and events described by Cao, then the way they constructed the author’s and characters’ nested thoughts and feelings is different from the way any other reader would
construct it, and it is lost to us. And when later-day scholars of *Stone* discuss Odd Tablet’s recorded response to a particular scandalous episode (a suicide of the main protagonist’s niece) and Cao’s subsequent incomplete revision of that episode, these scholars construct nested thoughts and feelings of the characters involved in that episode differently from lay readers.

(11) Or—to dip into a different interpretive tradition—when the nineteenth-century commentator Zhang Xinzhi writes that the “entire text of *The Story of the Stone* can be summed up in one phrase from the [fourth century BC *Commentary of Zuo*], ‘condemnation for failure to instruct,’” he constructs a nesting involving Cao’s intentions, mental states of characters from *Stone*, and his own attitude toward the events described in ancient historical narrative (which themselves feature a complex nesting of mental states).

So, on the one hand, any interpretation of *Stone* that aspires to be plausible and nonreductive must function on a high level of sociocognitive complexity. On the other hand, at any given point in the novel, Odd Tablet’s sociocognitive complexity is not Zhang Xinzhi’s; a native Chinese speaker’s sociocognitive complexity is not the same as someone’s who reads *Stone* in translation; and my students’ sociocognitive complexity is not mine. To speak the language of a novel (i.e., to explain to someone who hasn’t read it what is special about it) we have to nest mental state within mental state within yet another mental state, but the configuration and content of such nestings will differ from one reader to another.

Some nested interpretations may gain wider acceptance among particular groups of readers in particular cultural milieus. Since I am about to offer a close reading of several passages from *Stone*, I hope that some readers will find my nestings compelling, yet I will judge the effectiveness of my argument along different lines. I will consider it effective if, while disagreeing with my interpretations, you notice that your alternative interpretations still nest mental state within mental state within yet another mental state.

The first part of my chapter (sections 1 through 4) explores contrasts between explicit and implied nested mental states in Cao’s novel. Here my goal is to see if a text as overanalyzed as *Stone* may open up in new ways if we focus squarely on how it nests mental states. In the second part (sections 5 and 6), I take a different approach, thinking further through the concept of sociocognitive complexity and using *Stone* merely as a case in point rather than a focal point of analysis. Thus section 5 considers the possibility of a computer program that would count mental states in fiction (a spoiler: I believe that such a program will fail, but fail selectively and hence instructively); while section 6 brings in research of cognitive scientists who study mental states in stories for young children. I conclude with a brief discussion of my decision to use *Stone*—as opposed to a work that may be more familiar to most readers of this volume—as a case study.

1. “I Think I Know What You Meant”: Explicit Mental States

All novels build on our theory of mind, but some novels also make thinking about thinking their overt theme. *The Story of the Stone* is one of those. Its characters spend most of their waking hours in other people’s heads. This leads to some spectacular instances of explicit discussions of nested mental states.
Many of those discussions originate with the *Stone’s* main protagonists: a boy named Jia Bao-yu and his cousin, Lin Dai-yu. Bao-yu is afflicted with the “lust of the mind” (意淫). He wants to understand and share the emotions of girls, dozens of them, servants, cousins, and young aunts, populating the Jia’s sprawling aristocratic households (an ambition hardly compatible with his position as the male heir on whom the family’s hopes of future prosperity are pinned).\(^{16}\)

Dai-yu, on the other hand, always worries about what other people think about the propriety of her behavior. Brilliant poet and astute observer, she uses her formidable intellect to plumb ever-new heights of social paranoia. Dai-yu and Bao-yu are in love, but instead of bringing them together, their intense emotional bond drives them apart. Dwelling on the beloved’s thoughts turns out to be grist for the mill of the fate that long decreed their separation.

Here are two typical examples of explicit nested mental states that involve Dai-yu’s overreading of others.

Dai-yu and Bao-yu are visiting their cousin, Xue Bao-chai (whom Dai-yu considers her rival for Bao-yu’s affections). As they are sitting there, chatting and drinking tea and wine, Dai-yu’s maid, prompted by another maid, brings her a hand-warmer, and Dai-yu scolds her for it. Neither Bao-yu nor Bao-chai say anything, though for different reasons. Bao-yu knows “perfectly well” that Dai-yu’s intricately-phrased rebuke was “really intended for him,” but he makes “no reply, beyond laughing good-humoredly,” whereas Bao-chai, “long accustomed to Dai-yu’s peculiar ways,” simply ignores her words. However, Bao-chai’s mother, Mrs. Xue, is deaf to such intricacies and takes Dai-yu’s complaint at its face value. She points out to Dai-yu that it was “nice” of Dai-yu’s maids to think of her, because she often feels chilly. Dai-yu responds thus:

> You don’t understand, Aunt... It doesn’t matter here, with you; but some people might be deeply offended at the sight of one of my maids rushing in with a hand-warmer. It’s as though I thought my hosts couldn’t supply one themselves if I needed it. Instead of saying how thoughtful the maid was, they would put it down to my arrogance and lack of breeding.\(^{17}\)

Dai-yu is *imagining* people who’d *think* that she *thinks* that they are not taking good care of her. That’s bad enough, yet Dai-yu apparently goes easy on her aunt, who, after all, can only respond with the head-scratching “you are altogether too sensitive, thinking of things like that... Such a thought would never have crossed my mind.”\(^{18}\) It gets worse when Dai-yu’s audience is Bao-yu alone. Then Dai-yu can really spread her wings. Bao-yu’s “lust of the mind”—that is, his sympathetic interest in girls’ feelings—makes him a particularly inviting audience for Dai-yu’s paranoid nestings.

At Bao-chai’s birthday party, while the family is watching a play performed by a group of professional child actors, her aunt, Wang Xi-feng, observes smilingly that the way the boy who plays the main heroine “is made-up makes him look so like... someone we know.” Bao-chai and Bao-yu merely nod without responding (they know better), but another young relative, Xiang-yun, is “tactless enough” to blurt out that the actor looks like Dai-yu. Bao-yu shoots “a quick glance in [Xiang-yun’s] direction; but [it’s] too late;”\(^{19}\) for now the other guests catch on to the resemblance and start laughing.

Shortly after the party breaks up, the offended Xiang-yun orders her maid to start packing. Bao-yu overhears it and attempts to make her change her mind, explaining that the only reason he gave her that look is that he “was worried for [her] sake.” He knew that Xiang-yun didn’t know how sensitive Dai-yu can be and “was afraid that [Dai-yu] would be offended with [Xiang-yun].” Xiang-yun won’t have any of it. To her, Bao-yu’s glance implied that he thinks that she is
“not in the same class” as Dai-yu and hence mustn’t make fun of “the young lady of the house.”

I condense their conversation here, but you can see even from this condensed version that it consists of a series of third-level nestings all involving Xiang-yun’s perception of Bao-yu’s intentions regarding Dai-yu feelings.

But then it turns out that Dai-yu overheard Bao-yu’s conversation with Xiang-yun, so the real fun begins. First Dai-yu “coldly” explains to Bao-yu that even though he didn’t compare her with the child actor and didn’t laugh when others did, his secret thoughts, of which she’s apparently the best judge, implicate him severely. In the long quote that follows, the italics are in the original:

“You would like to have made the comparison; you would like to have laughed,” said Dai-yu. “To me your way of not comparing and not laughing was worse than the others’ laughing and comparing!”

Bao-yu found this unanswerable.

“However,” Dai-yu went on, “that I could forgive. But what about that look you gave Yun? Just what did you mean by that? I think I know what you meant. You meant to warn her that she would cheapen herself by joking with me as an equal. Because she’s an Honourable and her uncle’s a marquis and I’m only the daughter of a commoner, she mustn’t risk joking with me, because it would be so degrading for her if I were to answer back. That’s what you meant, isn’t it? Oh yes, you had the kindest intentions. Only unfortunately she didn’t want your kind intentions and got angry with you in spite of them. So you tried to make it up with her at my expense, by telling her how touchy I am and how easily I get upset. You were afraid she might offend me, were you? As if it were any business of yours whether she offended me or not, or whether or not I got angry with her!”

Bao-yu responds to this by becoming dejected at his incapacity to translate his “good intentions” into effective communication. Then he writes a poem “in imitation of a Buddhist gāthā,” wishing for the ability to know and convey one’s feelings without words. Then he “fear[s] that someone reading his gāthā might not be able to share his enlightenment” and adds “another set of verses after it to explain his point.” When Dai-yu later comes across the gāthā and the second poem, she adds to it another gāthā that parodies Bao-yu’s praise of wordless communications. And so it goes on—nesting upon nesting of involved discussions of feelings and of the futility of involved discussions of feelings.

(p. 182) 2. Implied Mental States

Bao-Yu and Dai-Yu’s private mind reading travails are but one instance of the malady afflicting all the Jias. The clan’s daily life is a complex network of social manipulation. Characters seek to anticipate and control the emotional responses of others, yet in the long run, their plans backfire. People do not respond as their would-be manipulators hoped they would. The craftiest mind readers, such as the beautiful and ambitious Wang Xi-Feng (who is in “charge of household management” because the male and female elders have abdicated their responsibilities), come to pitiful ends. Striving to read minds yet lacking self-awareness, the clan crumbles.
Here’s a typical Xi-feng moment, involving her philandering husband, Jia Lian, and her trusted maid, Patience. One day, as Xi-feng and Jia Lian are talking together about Jia Lian’s recent long trip, they hear voices in the next room. When Xi-feng asks who it is, Patience comes in to explain that “Mrs. Xue sent Caltrop [her maid and her son’s “chamber-wife,” i.e., concubine] over to ask [Patience] about something,” and that Caltrop has already received her answer and is gone. “Apparently pleased” by the mention of Caltrop, Jia Lian recollects that he saw her earlier that day and that she looks “most attractive.” Xi-feng then suggests that if Jia Lian likes Caltrop, Xi-feng will exchange Patience for her, so that Caltrop will become Xi-feng’s new maid and Jia Lian’s chamber-wife.

At this point, Jia Lian is called away, and, once he leaves the room, Xi-feng asks Patience “what on earth did Mrs. Xue want, sending Caltrop here like that.”

“It wasn’t Caltrop!” said Patience. “I had to make something up and hers was the first name that came to mind. [A woman who owes Xi-feng money came over to pay the interest.] It’s lucky I was in the outside room when she came, otherwise she might have blundered in here and Master would have heard the message. And we all know what Master is like where money is concerned. . . . Once he found out that you had savings, he’d pluck up courage to spend them in no time. Anyway, I took the money from her quickly and gave her a piece of my mind—which I am afraid you must have heard. That’s why I had to say what I did. I’d never have mentioned Caltrop in the Master’s presence otherwise!”

Xi-feng laughed.

“I was going to say! Why, for no apparent reason, should Mrs. Xue choose a chamber-wife to send here the moment Master gets back? So it was you up to your tricks, you little monkey!”

Here Jia Lian comes back and the husband and wife resume their conversation, but neither mentions Caltrop again.

If you look for references to mental states, you notice that Jia Lian thinks that Caltrop is attractive and that he likes money; that Xi-feng is surprised that Mrs. Xue would send in Caltrop while Jia Lian is in; and that Xi-feng is willing to get a new chamber-wife for her husband. These are explicit mentions of thoughts, feelings, and attitudes, and they describe isolated mental states. The meaning of the scene, however, resides not with (p. 183) them, but with the complex nestings that are implied rather than explicitly referred to by the narrator.

For instance, why is it that neither Xi-feng nor Jia Lian returns to their discussion of Caltrop once Jia Lian comes back? It must mean that Jia Lian has known all along that the madly jealous Xi-feng would never allow him to bring in Caltrop as a chamber-wife, and that she was merely playing with him, pretending to be a dutiful wife who wants her husband to have a new concubine and hence another shot at a son. Moreover, Xi-feng knows that Jia Lian knows that Xi-Feng merely pretends to be magnanimous about a concubine, just like she knows that he knows that she would never want to part with Patience, who is smart and loyal, more a friend that a servant.

In fact, we have just seen the amazing Patience in action. Because she knows that Jia Lian shouldn’t find out about the money, she figures that the best way to distract him is to make him think about a pretty girl. She also knows that if Jia Lian starts thinking about the girl, Xi-feng will be unhappy, yet that she will be more unhappy if Jia Lian finds out about the money. Patience counts on Xi-feng’s appreciating her “Caltrop” ruse upon finding out that it was meant to protect her purse, and she knows that Xi-feng will be able to use her husband’s lustful musings about Caltrop to remind
him who’s really in charge in their family.

At the same time, Xi-feng’s talk about trading Patience for Caltrop has a certain edge. It is as if Xi-feng were reminding Patience that she could exchange her for another maid if she wanted. Because Xi-feng doesn’t know what happened in the next room and suspects that Mrs. Xue would not send her son’s beautiful concubine with a message to Xi-feng at the time when Jia Lian is sure to be around, she knows that something is up with Patience’s mention of Caltrop. We thus can only guess if she is playfully teasing Patience when she proposes to Jia Lian, in Patience’s hearing, to trade her for Caltrop, or if she is quietly warning her that her position of trusted confidante is only as secure as Patience’s latest demonstration of absolute loyalty.

Though not a stupid man, Jia Lian is always at least one step behind both Xi-feng and Patience in their mind games. His relative cluelessness is consistent with the pattern we find elsewhere in the novel. Cao correlates his characters’ sociocognitive complexity with their age, gender, and class. That is, his young women of any class standing are much more likely to be capable of contemplating complex nested mental states than are rich men (such as Jia Lian) and older rich women (such as Mrs. Xue).

In fact, that’s yet another insight made possible by the cognitive perspective: we’ve always known that the author’s sympathies lay with his young female characters, but now we see just how he makes his young women sympathetic. Not that readers automatically sympathize with any character who is more sociocognitively complex than others. While such a character comes across as more interesting, she may also seem unpleasantly Machiavellian. To remain appealing, she has to be somewhat of an underdog. Think of Austen’s Fanny Price, who at any given point is more likely to be aware of other people’s feelings about other people’s feelings than anyone else in Mansfield Park, but who is also downtrodden, “timid, and exceedingly nervous.” Similarly, Cao keeps his highly sociocognitively complex young women sympathetic by making them sick, powerless, or doomed.

(Wang Xi-feng is an interesting case in point. She comes across as fun but also manipulative and dangerous. When I teach Stone to undergraduates, she is the one they hate to love. It’s open to debate to what extent her character is “redeemed” by marital unhappiness, illness, and eventual sad demise.

3. Gaps and Groans

When there is a gap between mental states explicitly mentioned in the text and mental states that we have to construct in order to make sense of what’s going on, no such construction is ever final. Take Xi-feng’s suggesting, in Patience’s hearing, that she should be traded for Caltrop. Xi-feng’s stated intention makes sense only if we simultaneously process a complex implied nesting that flatly contradicts that stated intention. But such nestings are always speculative. For instance, I just offered two possible interpretations of Xi-feng’s behavior. I don’t know which one of them is one is correct—or if both are correct—or if there is another, equally convincing interpretation.

What’s important is that, if a speculative interpretation is to be plausible, it must exhibit high sociocognitive complexity. The extremely involved social situation created by Cao cannot be understood and appreciated without nesting a mental state within a mental state within another mental state. “Xi-feng wants to trade Patience for Caltrop”...
captures exactly nothing. “Xi-feng wants to scare Patience” captures little. “Xi-feng wants Patience to be amused about the cat-and-mouse game she’s playing with her husband” or “Xi-feng wants Patience to remember that she would not forgive her a disloyalty” begin to get there.

Observe the contrast with scenes that are equally sociocognitively complex, but in which the gap between explicitly described and implied mental states is minimal, as it is in the two earlier episodes involving Dai-yu. My students groan in exasperation—and so do I—as we attempt to disentangle a Dai-yu argument. By and large, however, what she says usually coincides with our own sociocognitive map of the scene. It’s as if there were nothing left for us to interpret. We are lucky just to be able to follow Dai-yu’s torturous reasoning about what she thought Bao-yu wanted Xiang-yun to think about Dai-yu.

Cao’s novel features both kinds of high sociocognitive complexity; between gaps and groans the story gets told.

4. Sociocognitive Complexity across Chapters

So far we focused on implied nestings limited to a single sentence or to a couple of adjacent paragraphs. But, of course, sociocognitive complexity works across chapters, too. (p.185) Local implied nestings enter into implied nestings that span the length of the whole novel.

Take again Dai-yu’s annoyed reference to her cousin as “this Xue Bao-chai.” Five hundred pages later, in chapter 32, Dai-you reflects that if she and Bao-yu were destined for each other, why “did there have to be a Bao-chai…” This is essentially the same phrase (“this Xue Bao-chai” and “a Bao-chai”: “一個薛寶釵” and “一寶釵”), and it reflects Dai-yu’s anguished sense of propriety. She can’t say anything harsh or vulgar, so a reference to Bao-chai preceded by the vaguely dismissive “this” or “a” becomes a signature expression of her irritation and unhappiness.

On both occasions, the phrase “this [Xue] Bao-chai” introduces a separation of Dai-yu and Bao-yu. In chapter 5, we first learn that, sharing Grandmother Jia’s quarters as children, Dai-yu and Bao-yu developed “an understanding so intense that it was almost as if they had grown into a single person.” Then “this Xue Bao-chai” appears on the scene, and it becomes clear that Dai-yu and Bao-yu don’t have a perfect understanding. Specifically, the narrator explains that Dai-yu was “put out” by Bao-chai’s popularity, but, as for Bao-yu, he remained unaware of the complexities of the situation and, moreover, that he and Dai-yu had “occasional tiffs and misunderstandings that are usual with people who have a great deal to do with each other.” Finally, we have a brief account of one such quarrel, with Dai-yu “crying alone in her room and Bao-yu feeling remorsefully that perhaps he [has] spoken too roughly” and going in “to make his peace with her,” after which, “gradually, very gradually, Dai-yu’s equanimity [is] restored.”

In chapter 32, we once more start with the image of Dai-yu and Bao-yu together—Bao-yu is viewed as Dai-yu’s “true friend,” or “soul mate”—and end with Dai-yu’s crying alone. Haun Saussy’s interpretation of this latter episode applies equally well to the one we just discussed. As he explains (operating, here and elsewhere, on at least the third level of sociocognitive complexity), “falling back from the eager hypothesis of a ‘we’ to the harsh fate of an ‘I’ leaves a dejected [Dai-yu] in tears, which the narrator uses as a way to bring the focus of the story back into the external world of visible actions and reactions.”
Still more sociocognitive complexity: Saussy considers this episode an example of free indirect discourse particularized through the context of Chinese literary history. As he puts it, “ambiguity about the source of narration riddles the whole scene. For the sequence purporting to convey [Dai-yu’s] intimate unspoken thoughts is quite literally a self-commentary on the narrator’s own wording, a gesture that draws attention to the medium in all its artifice.”

The repetition of “this [Xue] Bao-chai” thus creates a dialogue between the two scenes. Becoming aware of this dialogue means constructing cross-novel implied nestings that involve the narrator and his implied readers. We may say, for instance, that the narrator wants to draw his readers’ attention to Dai-yu’s tortured commitment to propriety even in the midst of anguish. Or we may say that the narrator wants us to be aware of the intentions of fate. Because Dai-yu and Bao-yu’s separation has been predestined, it’s only fit that most of their interactions must end in her tears and his disappointment. “This Xue Bao-chai” is but a tool used by fate—but also by the narrator.

5. Can a Computer Program Count Nested Mental States in *The Story of the Stone*?

If, as I claim, works of fiction always function on at least the third level of sociocognitive complexity, can one design a computer program that will count mental states in a given sentence, paragraph, or chapter? The possibility of such a program has been mentioned to me on several occasions, with cautious enthusiasm by cognitive scientists and computer scientists and with dread by my colleagues from literary studies. I would love to see a computer trying to count mental states in fiction because I believe that its failure would be as illuminating as was the failure of various artificial intelligence projects in the 1950s–1970s.

The latter, as you may remember, alerted scientists to the untold complexity of evolved human cognition. The machines could not replicate cognitive processes that came so easily to people that they hadn’t even been aware of them. Just so, by failing to register nested mental states in fiction, a computer program would illuminate cognitive processes that make reading fiction possible and that we take completely for granted, such as a constant attribution of nested mental states to characters, readers, (implied) authors, and narrators.

What will be particularly instructive in this case is that the failure might be selective. I believe that a computer would be able to count mental states in many works of mainstream popular fiction, but would fail miserably with more complex texts. The reason for that is that works of popular fiction create high sociocognitive complexity by nesting mental states of their characters and by describing these mental states explicitly. A computer program can count those. In contrast, works of fiction that we call literature and tend to teach in college, nest, to a much greater degree, mental states of narrators and implied authors and readers. They also make us work harder at figuring out mental states of both characters and narrators because they often imply, rather than explicitly name these mental states.

Consider this passage from John Irving’s novel *The 158-Pound Marriage* (1974):

“I am going to get a lover,” she said, “and I’m going to let you know about it. I want you to be embarrassed when you make love to me wondering if I am bored, if he does it better. I want you to imagine what I say that I can’t say to you, and what he has to say that you don’t know.”
You can indeed design a computer program that would count mental states in such a novel. Make it pick such words as “want,” “embarrassed,” “wonder,” “bored,” “can’t say,” “imagine,” and you will have a fairly accurate map of a given sentence’s sociocognitive complexity. “I want you to be embarrassed because you wonder if I am bored”—that’s fourth-level sociocognitive complexity, and a computer may just be able to perform this calculation.

(p. 187) In contrast, faced with “And now suddenly this Xue Bao-chai had appeared on the scene—a young lady who, though very little older than Dai-yu, possessed a grown-up beauty and aplomb in which all agreed Dai-yu was her inferior,” a computer will have nothing to go by except the word “agreed” (贊). But, as we’ve already seen, that word contributes little to the sociocognitive complexity of that sentence. A computer program can’t register implied mental states, much less figure out context-specific relationships that organize these mental states into nestings. Because in The Story of the Stone, any word—including “a” and “this”—can create an implied nested mental state, only a human mind, with its infinite sensitivity to contexts, can follow it.

But what about such passages from Stone that nest mental states of characters and spell those out? After all, Dai-yu’s speech about the look that Bao-yu gave to Xiang-yun (i.e., “I think I know what you meant”) is not terribly different from Irving’s “I want you to be embarrassed when you make love to me wondering if I am bored.” It seems that Cao uses mixed techniques for creating sociocognitive complexity, that is, that he uses both the techniques that we may associate with “more sophisticated” works of fiction (i.e., implied mental states of not just characters but also of the narrator, the implied author, and the implied reader) and those that we may associate with mainstream, popular fiction (i.e., explicit descriptions of characters’ mental states).

In fact, one wonders if this may not contribute to the peculiar dual position of The Story of the Stone in Chinese culture. As Dore J. Levy puts it, to appreciate this position, “we must imagine a work with a critical cachet of James Joyce’s Ulysses and the popular appeal of Margaret Mitchell’s Gone with the Wind.” Of course, the mixed pattern of creating nested mental states is by no means a sufficient explanation of Stone’s iconic status. Still, it suggestively complements other critical explanations that grapple with various dualities at the heart of the novel and with its simultaneous appeal both to literary cognoscenti and to popular audiences.

Cognoscenti, though, may have the last laugh because even the scenes that seem straightforward in their explicit attribution of mental states to characters can be read in terms of implied mental states of the narrator, the implied reader, and the implied author. We only need to remember that Dai-yu’s explicit nestings are framed by the narrative as means to a very particular end (in contrast, for instance, to the straightforward nestings in Irving’s The 158-Pound Marriage, which are ends in and of themselves).

Worrying about what others are thinking makes Dai-yu anxious and distressed. And she must be kept anxious and distressed because it enables her to pay to Bao-yu the “debt of tears.” Dai-yu incurred this debt in her previous existence as the “Crimson Pearl Flower,” whom Bao-yu—back then the “Divine Luminescent Stone-in-Waiting”—watered “daily with sweet dew, thereby conferring on her the gift of life.” Neither Dai-yu nor Bao-yu know about their past lives, but the reader is free, or encouraged to contemplate the intentions of fate (and of its various stand-ins) every time Dai-yu and Bao-yu interact with each other.

This is yet another aspect of the novel that a computer program counting mental states will miss. It may count explicit references to thoughts and feelings in Dai-yu’s tirades, but it won’t be able to register the layering of mental states that
occurs when the reader notices herself stopping in her tracks and asking if it’s been intended all along that Dai-yu must overthink Bao-yu’s thinking about Xiang-yun’s thinking.

That reader may ultimately decide, together with Anthony C. Yü, that the solitary, enigmatic girl in the narrative becomes a person who has so powerfully laid claim to the care and concern of readers down the ages that no aetiological myth of predestined suffering can remove the authenticity stamped on her private grief and public despair.42

Yet note how this conscious resistance of the reader to the “aetiological myth of predestined suffering” is itself a complex nesting of implied mental states. We can map it, very crudely, along the lines of, I refuse to think that when Dai-yu overthinks Bao-yu’s thinking about Xiang-yun’s thinking, she is being but a toy in the hands of destiny that drives her to her preappointed end. Given Stone’s framing as a meditation on predestination and memory, no explicit nestings of its mental states remain immune to being wrapped in implied nestings—and then rewrapped in implied nestings that resist the original implied nestings.

6. What Rosie Knew

How early does it start? Would computers be able to count mental states in stories for very young children? As it turns out, if they would, you wouldn’t want to read such stories to your children.

In 2004, developmental psychologists Joan Peskin and Janet Wilde Astington decided to explore further the connection between the acquisition of vocabulary in young children and development of theory of mind.44 It’s been shown that children attending schools in low-income neighborhoods “demonstrate substantial lags in their theory-of-mind understanding” and that at six years old, they know only half the number of words as do children from higher socioeconomic groups:

Children whose parents do not provide a rich lexicon for distinguishing language about perceiving, thinking, and evaluating might make important gains from hearing and talking such talk in their everyday story reading. . . . A rich vocabulary, more than any other measure, is related to school performance.45

Peskin and Astington wanted to test whether exposure to an explicit discussion of mental states (they call it metalanguage) “will result in a greater conceptual understanding of one’s own and other people’s beliefs or whether this understanding develops more implicitly.”46 They rewrote kindergartners’ picture books “specially for the study so that the texts were rich in explicit metacognitive vocabulary, such as think, know, remember, wonder, figure out, and guess, in both the texts and text questions.”47

Thus Pat Hutchins’s classic Rosie’s Walk (1968)—which features a chicken on her daily walk, unaware that a hungry fox is right behind it—was altered to include such descriptions of the chicken’s thoughts as, “does Rosie know that Fox has been following her? No, Rosie doesn’t know. She doesn’t even guess.” The children in this “explicit metacognitive condition were compared with a control group that received the identical picture books, with a similar number of words and questions, but not a single instance of metacognitive vocabulary.”
What Peskin and Astington found was that “hearing numerous metacognitive terms in stories is less important than having to actively construct one’s own mentalistic interpretations from illustrations and text that implicitly draw attention to mental states.” Children exposed to explicit metacognitive terms did start using them more, but they used them incorrectly.

These results support earlier studies, one of which found that “children exposed to more metacognitive terms of certainty (think, know, and guess) in a television show later displayed a poorer understanding of certainty distinctions than those exposed to episodes containing fewer of these terms.” Two other studies, “which compared children whose teachers used more metacognitive vocabulary to those whose teachers used less, found superior performance on theory-of-mind tasks for children whose teachers used fewer metacognitive terms.”

To explain such counterintuitive findings, Peskin and Astington suggest that “the teaching of information does not automatically lead to learning.” What is required instead is a “constructive, effortful process where the learner actively reorganizes perceptions and makes inferences. . . . These inferences lead to an understanding that may be all the deeper because the children had to strive to infer meaning. Ironically, the more direct, explicit condition may have produced less conceptual development precisely because it was explicit.”

What do Peskin and Astington recommend for fostering constructive learning? Reading fiction.

Dramatic tension in stories is created when the various characters have disparate knowledge with regard to the action. This may be through error: The reader knows that Romeo does not know that Juliet lies drugged, not dead. Or it may be through deception: Pretending his assigned chore is an adventure, Tom Sawyer tricks his friends into whitewashing the fence.

The examples chosen by Peskin and Astington are chock-full of implied third-level巢ings. To stay just with the action that they describe above (and thus ignoring, for instance, complex, and, perhaps, more interesting, implied nestings created by the tone of Twain’s narrator), Tom didn’t want his friends to realize that he hated whitewashing the fence. Just so, Romeo didn’t know that Juliet wanted some people to think that she is dead. Neither Shakespeare nor Twain spells out those mental states for his readers; we have to deduce them ourselves in order to make sense of what we read.

Think about it. Works of fiction that do not spell out nested mental states may enrich understanding of mental states, foster the ability for constructive learning, and improve vocabulary in preschool and school-age children. I wouldn’t claim that the effect is exactly the same for grown-ups. After all, theory of mind goes through some major developmental milestones in young children and adolescents, so the impact might be more pronounced for those age groups. (Nor am I discussing here the proximate effect of reading fiction at any age: pleasure.) Still, it seems to me that the difference between fictional narratives that require readers to work harder to deduce implied nested mental states, on the one hand, and fictional narratives that spell out mental states, on the other, emerges as a fascinating and underexplored topic. As such it certainly warrants a closer look from literary scholars (particularly those who are concerned about the diminishing role of the humanities and wish to argue for the importance of sustained exposure to challenging literary texts on every educational level).

Peskin and Astington’s examples are firmly ensconced in Western literary tradition and as such don’t need much
context. Most readers of this volume will immediately recognize the relevant scenes from *Romeo and Juliet* and *Tom Sawyer*. In contrast, some of them may not have heard of *The Story of the Stone*, for it is, alas, still not as widely known as it deserves to be. Thus, as I conclude my chapter, something must be said about my decision to try out a new idea—nested mental states as units of meaning in fiction—on a novel that is itself new to some readers.

I am thinking of Haun Saussy’s comment on something that David Hawkes, the novel’s translator, wrote in 1973. In the introduction to the first volume of *Stone*, Hawkes offered a brief yet compelling Freudian reading of Bao-yu’s behavior on a particular occasion.55 Here is Saussy reflecting on it in 2012:

The gesture is double: it advances a claim both for the importance of the novel (that it anticipated Freud; that it thereby resembles other masterworks, by Sophocles, Shakespeare, Leonardo da Vinci; that it offers insights into universal human nature) and for the importance of the interpretation (that Freudian theory applies not only to Western literary works but also to works from [different] times and places).56

I can’t say it better, so I will adapt it. I chose Cao’s novel as my case study because it brilliantly anticipated cognitive literary theory. I also wanted to show that cognitive literary theory applies not only to Western literary works but also to those from different times and places. (In fact, the roots of cognitive literary theory are in comparative and world literature.)57 So if you want to see how the social becomes the literary and how the literary plays games of hide-and-seek with mental states, read *The Story of the Stone*.

**Acknowledgments**

I thank J. Keith Vincent for introducing me to *Honglou meng* and David L. Rolston for his thoughtful and detailed suggestions (some of which I couldn’t fully implement, alas, due to length limitations). I am also grateful to Elaine Auyoung, Ellen Spolsky, and Paul L. Harris for their helpful comments.

**Works Cited**


Auyoung, Elaine. Email communication, February 7, 2014.


Find this resource:


Find this resource:


Find this resource:


Find this resource:

———. Email communication, April 18, 2014.

Find this resource:


Find this resource:


Find this resource:


Find this resource:


Find this resource:


Find this resource:


Find this resource:

  
  Find this resource:

  
  Find this resource:

  
  Find this resource:

  
  Find this resource:

  
  Find this resource:

———. Email communication, June 6, 2014.
  
  Find this resource:

  
  Find this resource:

  
  Find this resource:

  
  Find this resource:

  
  Find this resource:

  
  Find this resource:

  
  Find this resource:

Find this resource:


Find this resource:


Find this resource:


Find this resource:


Find this resource:


Find this resource:


Find this resource:


Find this resource:

**Notes:**


(2). I am indebted to Elaine Auyoung for reminding me how crucially emotions in general and social anxiety in particular are implicated in our nesting of mental states.


(5). Note that in the original, even that mental state is not present in this explicit form. The word 調 implies a verbal agreement rather than a mental state. 多謂 is “all said” rather than “all agreed”—although this is a situation in which the boundary between the two is blurry.


(9) A cognitive perspective explored in this study thus builds on the classic argument of narrative theory, that in “any example of narrative art there are, broadly speaking, three points of view—those of the characters, the narrators, and the audience. As narrative becomes more sophisticated, a fourth point of view is added by the development of a clear distinction between the narrator and the author” (Robert Scholes, James Phelan, and Robert Kellogg, The Nature of Narrative: Revised and Expanded, 40th anniversary edition [New York: Oxford University Press, 2006], 240). For a pioneering application of the study of narrative irony, emerging from the distance among the “four different points of view, namely those of the character, the narrator, the reader, and the implied reader” to The Story of The Stone, see Wong Kam-Ming, “Point of View, Norms, and Structure: Hung-Lou Meng and Lyrical Fiction,” in Andrew H. Plaks, ed., Chinese Narrative (Princeton: Princeton University Press, 1977), 203–10.


(12) See Wai-Yee Li’s discussion of Red Inkstone’s invoking an “inner circle” of people who “know” how to read particular parts of Honglou meng (Enchantment and Disenchantment: Love and Illusion in Chinese Literature [Princeton: Princeton University Press: 1993], 198.)


(14) In China, Stone is both a revered masterpiece and a living source of daily cultural references. The profusion of publications on the novel has led to the creation of a dedicated scholarly discipline, “Redology” (Hongxue 紅學). Those less academically inclined can visit Beijing and Shanghai theme parks based on the novel, argue the relative merits of its numerous televised versions, and buy “playing cards, teapots, CDs, stamps, comic books, pottery, snuff bottles, lanterns, vases, figurines, coins, and ashtrays, merchandise of every quality,” featuring its characters and settings (Andrew Schonebaum, “Introduction,” in Andrew Schonebaum and Tina Lu, eds., Approaches to Teaching “The Story of the Stone” [New York: Modern Language Association, 2012], 6). The popular feeds back into the scholarly: “The Journal of Stone Studies devoted more than 150 pages” of one of its issues to the 1987 television

(15) . See Zunshine, Why We Read Fiction: Theory of Mind and the Novel (Columbus: Ohio State University Press, 2006).

(16) . In David L. Rolston’s apt characterization, Bao-yu is “less the central actor than the reflector or center of consciousness of the narrative” (Traditional Chinese Fiction and Fiction Commentary: Reading and Writing Between the Lines [Stanford: Stanford University Press, 1997], 197).


(18) . Ibid.


(20) . Cao, The Story of the Stone, 1:437.


(22) . Cao, The Story of the Stone, 1:442.

(23) . Dore J. Levy, Ideal and Actual in “The Story of the Stone” (New York: Columbia University Press, 1999), 47. For an invaluable analysis of Jia family dynamics, see Levy’s chapter “Family Togetherness.”


(26) . Xi-feng wouldn’t want to bring in a new chamber-wife for Jia Lian even if, as David Rolston reminds me, “the reader would also assume that as with the case of Patience, Xi-feng would do her best to make sure she is chamber-wife in name only” (email communication, June 6, 2014).

(27) . Jia Lian and Xi-feng only have a daughter, and Xi-feng doesn’t seem to be able to have more children (later in the novel, she has a miscarriage).

(28) . Evil masterminds tend to be highly sociocognitively complex. Compare to Blakey Vermeule’s important discussion of masterminds in Why Do We Care about Literary Characters? (Baltimore: Johns Hopkins University Press, 2010), 86. See also my discussion of students’ perception of Mrs. Roby from Wharton’s “Xingu” in “‘Theory of Mind’ as a Pedagogical Tool.”


(31) . Cao, The Story of the Stone, 2.32.132. For the purposes of this discussion, I am omitting Dai-yu’s thoughts about
“gold and jade” and Bao-chai’s “golden locket.”


(34) . Cao, The Story of the Stone, 1:125.

(35) . It’s “true friend” in Hawkes’s translation and “soul mate” in Saussy’s “Unspoken Sentences,” 432.


(37) . Ibid., 433.


(40) . Levy, Ideal and Actual, 1.

(41) . Cao, The Story of the Stone, 1:53.


(46) . Ibid., 254.

(47) . Ibid., 255.

(48) . Ibid., 253.

(49) . Ibid., 265.

(50) . Ibid., 266.

(51) . Ibid., 267.
Of course, “the exact interpretation of [Peskin and Astington’s] results needs more research” (Paul L. Harris, email communication, April 18, 2014). To begin with, the emphasis on the importance of reading fictional stories that make children work hard at deducing mental states does not mean to downplay the crucial role of talking to children about thoughts and feelings, and it may shed an interesting light on the underlying structure of those conversations. See, for instance, the recent study by Harris and his colleagues, who looked at children’s attribution of emotions, attendant upon their attributions of false beliefs, and found that while the four- to six-year-olds may judge correctly that Red Riding Hood doesn’t know that the Wolf is waiting for her in her grandmother’s cottage, they may still say that she is afraid rather than happy as she approaches the cottage. While thus positing a lag between “children’s understanding of a protagonist’s mistaken beliefs and their grasp of the emotions that flow from such beliefs,” this study also found that “children with mothers who use more mental-state language make more correct attributions” of emotions (Harris et al., “The Mysterious Emotional Life of Little Red Riding Hood,” in Kristin H. Lagattuta, ed., Children and Emotion: New Insights into Developmental Affective Sciences [Basel: Karger, 2014] ). Moreover, as Harris observes elsewhere, “a simple count of mental-state terms [used by mothers] may not be the most sensitive measure of effective maternal input even if it is a useful correlate. [It’s possible] that is is the mother’s pragmatic intent, notably her efforts to introduce varying points of view into a given conversation, that is the underlying and effective source of variation” (“Conversation, Pretense, and Theory of Mind,” in Astington and Baird, eds., Why Language Matters, 77.)

I talk about pleasure at length in my Why We Read Fiction, even though that pleasure still falls into a suspect less pure category of pleasure, described in a conversation reported by Michael Bérubé, as “a form of pleasure that is good for you, and good for your brain” (“How We Got Here,” PMLA 128.3 [2013]: 537).

It would be interesting to see, for instance, how this difference maps onto works of traditional Chinese fiction. A tantalizing hint that this difference might have been present and theorized as a pedagogical strategy emerges from the work of Rolston, who points out that although “use of direct psychological description in fiction increased throughout the Ming and Qing dynasties, it was never popular or influential.” One important reason for it was that “the main justification for reading [fiction was] to develop the ability to judge human character; easy access to the inner life of characters would defeat this pedagogical purpose.” As traditional commentators saw it, “the author who is presented as the most subtle in his laying down of [ . . . ] clues that raise suspicions about a gap between an inner state of the character’s mind and his or her actions or words [ . . . ] becomes the author most worthy of praise” (Traditional Chinese Fiction, 217).


Lisa Zunshine

Lisa Zunshine is Bush-Holbrook Professor of English at the University of Kentucky and a former Guggenheim Fellow. She is the author or editor of eleven books, including Bastards and Foundlings: Illegitimacy in...