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 did¹). It seems to me that we do it often, though

¹ Notes:
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.

Fiction is where it gets interesting. Nested mental states 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*, also known as *Dream of the Red Chamber* (*Hong-Lou 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.”

---

I am grateful to J. Keith Vincent for introducing me to *Hong-lou Meng* and to key scholarship on the novel.

1 On the history, limitations, recent return, and perspectives of introspectionism in metacognition, see John Dunlosky and Janet Metcalfe, *Metacognition* (Los Angeles: Sage, 2009), 9-34.

今忽然來了一個薛寶釵，年歲雖大不多，然品格端方，容貌豐美，人多謂黛玉所不及。3)

What’s going on in this sentence? The narrator wants his readers to realize that, whereas Dai-yu, insecure as she is, is certain that everyone around her considers her inferior to Bao-chai, she might be misinterpreting their feelings. That’s at least four recursively nested mental states, but to articulate them, we have to take in subtle cues, such as the irritated 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.”4 Here is what’s new about the cognitive approach that I propose here. I focus on nested mental states as units of

---

Accessed September 6, 2013.

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.⁵ Consider theory of mind, that is, our evolved cognitive capacity for attributing mental states to ourselves and to others. 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.⁶ 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 non-fictional 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

7 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 Scholes, Phelan, and Kellog’s 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 *Chinese Narrative*, ed. Andrew H. Plaks [Princeton: Princeton UP, 1977], 203-210).
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. So does in fiction, understanding what we read in terms of our own, the characters,’ and the (implied) author’s thoughts and feelings, inevitably involve 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*
speculate about Odd Tablet’s presumed response to a particular scandalous episode (a
suicide of one of the main protagonist’s aunts) 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.

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 [4th 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
non-reductive 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.

---

*Zhang Xinzhi [Chang Hsin-chih], “How to Read the Dream of the Red Chamber,”
translated and annotated by Andrew H. Plaks; additional annotation by David L. Rolston.

University Press, 1990. 326. See also note 15.*
By way of analogy consider that in a group of people who speak the same language no two people will ever form the same exact sentences. Just so, to speak the language of the novel—any novel—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 essay (sections one through four) 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 Stone¹⁰ may open up in new ways if we focus squarely on how it nests

¹⁰ 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” (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 Approaches to Teaching The Story of the Stone, eds. Andrew Schonebaum and Tina Lu [New York: The Modern Language Association, 2012], 6). The popular feeds back into
mental states. In the second part (sections five and six), 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 five 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 six 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\(^1\), 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).

Dai-yu, on the other hand, always worries about what other people are thinking 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. Used to Dai-yu’s “peculiar ways,” neither Bao-yu nor Bao-chai say anything, but Bao-chai’s mother, Mrs. Xue, protests 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.\textsuperscript{12}

Dai-yu is \textit{imagining} people who’d \textit{think} that she \textit{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.”\textsuperscript{13} It gets worse when Dai-yu’s audience is Bao-yu, with whom 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, comments slyly on the resemblance between “someone we know” and a beautifully made-up boy who plays the main heroine. Bao-chai and Bao-yu merely nod without responding (once again, 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,”\textsuperscript{14} 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

\textsuperscript{12} Cao, \textit{The Story of the Stone}, 1:193.

\textsuperscript{13} Ibid.

\textsuperscript{14} Cao, \textit{The Story of the Stone}, 1:436.
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 everyone 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.”¹⁵

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

¹⁵ Cao, The Story of the Stone, 1:437.
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 short meditative poem wishing for the ability to know and convey one’s feelings without words. Then he “fear[s] that someone reading his [poem] might not be able to share his enlightenment” and adds “another set of verses after it to explain his point.” When Dai-yu finds the poem later, she adds to it two stanzas which parody 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.

2. Implied mental states

Bao-Yu and Dai-Yu’s private mindreading travails are but one instance of the malady afflicting all the Jias. The clan’s daily life is a complex network of social

---


manipulation. Every character worth her salt seeks 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 mindreaders, 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\(^\text{18}\)), come to pitiful ends. Striving to read minds but failing to see the writing on the wall, 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.”\(^\text{19}\) 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.”

---


\(^\text{19}\) Cao, *The Story of the Stone*, 1:308.
‘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 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 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 return 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 another 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

---

21 Jia Lian and Xi-feng only have a daughter, and Xi-feng doesn't seem to be able to have more children.
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.\textsuperscript{22} To remain appealing, she

\textsuperscript{22} Evil masterminds tend to be highly sociocognitively complex. Compare to Blakey Vermeule’s important discussion of masterminds in \textit{Why Do We Care About Literary Characters?} (Baltimore: The Johns Hopkins University Press, 2010), 86. See also my
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 contradictions

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. For instance, I just offered two possible interpretations of Xi-feng’s suggesting, in Patience’s hearing, that she should be traded for Caltrop. 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 an 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.

My next example of an implied nesting also centers on Dai-yu, but in this case, we have a large gap between implied and explicitly mentioned mental states. The famous fifth chapter of the novel (famous because it contains Bao-yu’s prophetic dream about the respective fates of the young women surrounding him) opens with an explanation of how dear Dai-yu has been to her Grandmother Jia, the clan’s matriarch. Since her arrival at the Jias,’ after her mother’s death, Dai-yu has received the same meals and accommodations
as Bao-yu—whom Grandmother Jia worships—and so has grown very close with Bao-yu:

[The] objects of [Grandmother’s] partiality themselves began to feel an affection for each other which far exceeded what they felt for any of the rest. Sharing each other’s company every minute of the day and sleeping in the same room at night, they developed an understanding so intense that it was almost as if they had grown into a single person.25

All this seemed to change, however, when another granddaughter, Bao-chai, came to visit the Jias. “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.”26

This is the same sentence that we looked at in the beginning of my essay. What I want to add now is that its sociocognitive mapping (i.e., “the narrator wants his readers to realize that, whereas Dai-yu, insecure as she is, is certain that everyone around her considers her inferior to Bao-chai, she might be misinterpreting their feelings”) depends on us seeing this as a case of free indirect discourse. Dai-yu’s private insecurity is presented as something that “all” (人多) agree on, a perception that can’t be correct,


26 Ibid.
because at least one person, Bao-yu, won’t be included in this snarky intermental unit, and neither, perhaps, will Bao-chai.

If we stay with this interpretation, we notice an important role that this sentence plays across the five opening paragraphs, widening the gap between explicitly mentioned and implied mental states.

We have just learned that, sharing Granmother Jia’s quarters as children, Dai-yu and Bao-yu, developed an intense understanding. We then hear Dai-yu’s anxious assertion that “all” consider her inferior to Bao-chai. The narrator then 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.”

But if Bao-yu does not share Dai-yu’s belief that Bao-chai makes Dai-yu look inferior, and if on top of that they have “tiffs” that make Dai-yu feel lonely, then perhaps, they don’t have “an understanding so intense that it was almost as if they had grown into

\[\text{27 For a discussion of intermental units, see Alan Palmer’s essay in this volume as well as Palmer, “Storyworlds and Groups,” in Introduction to Cognitive Cultural Studies, ed. Lisa Zunshine (Baltimore: The Johns Hopkins University Press, 2010), 184-186.}\]

\[\text{28 Cao, The Story of the Stone, 1:125.}\]
a single person.” Free indirect discourse thus introduces the first note of discord into Bao-yu and Dai-yu’s intermental unit; their quarrels tear that unit further asunder. The intentions of the narrator become inextricable from the children’s frustrated expectations of perfect understanding. We can say, for instance, that the narrator wants us to note how Dai-yu’s private certainty about everyone’s preferring Bao-chai, pulls the children apart; or that the narrator wants us to register the children’s pain at their realization that they don’t think alike.

You may come up with a different interpretation. But however you chose to explain what’s going on, to do justice to the complexity of the situation conjured by Cao, you have to operate on at least the third level of sociocognitive complexity. Moreover, you have to do it by evoking implied mental states which would differ from or even directly contradict the explicitly stated mental states, such as that the children have an intense understanding, or, that everyone agrees that Dai-yu is inferior to Bao-chai.

We’ve seen this before. Wang Xi-feng’s stated intention to trade Patience for Caltrop made sense only if we simultaneously processed a complex implied nesting that flatly contradicted that stated intention. We’ll see it again.

---

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. 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 thirty-two, 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 “一個寶釵”), 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 separation of Dai-yu and Bao-yu. Here, as in chapter five, we start with the image of Dai-yu and Bao-yu together—Bao-yu being Dai-yu’s “true friend,” or “soul mate”—and end with Dai-yu’s crying alone. As Haun Saussy 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

---


31 It’s “true friend” in Hawkes’s translation and “soul mate” in Saussy’s “Unspoken Sentences,” 432.
as a way to bring the focus of the story back into the external world of visible actions and reactions.”

Significantly, 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 (implicated, perhaps, in free indirect discourse). 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 1950-1970s.

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.

In contrast, faced with “And now suddenly this Xue Bao-chai had appeared on the scene . . .” a computer will have nothing to go by except the word “agreed” (謂). 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 complements suggestively 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 at will 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. Yu, 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.\textsuperscript{38}

Yet note how this conscious resistance of the reader to the interpretation expected of the implied reader is itself a complex nesting of implied mental states. We can map it, very crudely, along the lines of, I \textit{refuse} to \textit{think} that when Dai-yu \textit{overthinks} Bao-yu’s \textit{thinking} about Xiang-yun’s \textit{thinking}, she is being but a toy in the hands of destiny that drives her to her pre-appointed end. Given \textit{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 re-wrapped in implied nestings that resist the original implied nestings.

6. What Rosie Knew\textsuperscript{39}

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 well-established connection between the acquisition of


\textsuperscript{39} Parts of this section have appeared in Zunshine, “Why Fiction Does It Better,” \textit{The Chronicle of Higher Education} December 9, 2013.
vocabulary in young children and development of theory of mind. It’s been shown that children attending schools in low-income neighborhoods “demonstrate substantial lags in their theory-of-mind understanding” and that at 6 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.40

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.”41 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.”42

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 nestings. 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 merely wanted some people to think that she is dead. Neither Shakespeare nor Twain spell out those mental states for their readers; we have to deduce them ourselves in order to make sense of what we read.

---

45 Peskin and Astington, “The effects of adding,” 266.
Think about it. Works of fiction that do not spell out 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.47) 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 an important and underexplored topic. As such it certainly warrants a closer look from literary scholars (particularly those concerned about the diminishing role of the Humanities).

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 essay, 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.

47 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 Berube, as “a form of pleasure that is good for you, and good for your brain” (“How We Got Here,” PMLA 128.3 [2013], 537).
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. 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).

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). So if you want to see how the social becomes the

---


50 See in particular Patrick Colm Hogan’s groundbreaking “Literary Universals,” *Poetics Today* 18:2 (Summer 1997) and “Toward a Cognitive Science of Poetics:
literary and how the literary plays games of hide-and-seek with mental states, read The Story of the Stone.

Works Cited:


-----. “Style Brings In Mental States.” *Style* 45.2 (2011): 349-356.

-----. “Why Fiction Does It Better.” The Chronicle of Higher Education December 9
(2013): B4-B5.