How Memories Become Literature

Lisa Zunshine

➡ *For additional information about this article*

How Memories Become Literature

Lisa Zunshine

Abstract
Cognitive science can help literary scholars formulate specific questions to be answered by archival research. This essay takes, as its starting point, embedded mental states (that is, mental states about mental states) and their role in generating literary subjectivity. It then follows the transformation of embedded mental states throughout several manuscripts of Christa Wolf's autobiographical novel, Patterns of Childhood (Kindheitsmuster, 1976), available at the Berlin Academy of Arts. The author shows that later versions of Patterns of Childhood have more complex embedments in the chapter describing the adolescent protagonist's relationship with her schoolteacher. This textual development is integral to the process whereby the presumably authentic memories of the past are constructed to fit the present needs of the person who is doing the remembering. Accompanying the three case studies of the manuscript revision is a discussion of theoretical and practical implications of this "cognitive-archival" approach to literature.

Lying and pretense offer a hope for redemption in Christa Wolf's *Patterns of Childhood* (1976), an autobiographical novel about growing up in Nazi Germany, not least because the occasions on which Nelly Jordan (the stand-in for Wolf's younger self) had lied to various power-wielding adults seems to have remained engraved in her memory. As such, these occasions are invaluable for the grown-up narrator, who is trying to reconstruct her childhood as an enthusiastic member of the girls' wing of the Hitler Youth and is continuously stumped by various kinds of amnesia. As Wolf puts it, "Where Nelly's participation was the deepest, where she showed devotion, where she gave of herself, all relevant details have been obliterated" (229). Faking devotion—and being conscious of doing so—seems to have preserved some small parts of Nelly inviolate and also made the later project of autobiography possible.

Lying as a form of private resistance, as a path to a heightened self-awareness, and as a pledge of memory—Wolf makes a compelling case for all three, and we don't need to argue with her about it. What we can do—and by this "we" I mean literary scholars who work on inte-

grating insights from cognitive science with more traditional methods of study—is to highlight yet another side of lying that makes it propitious for writers. Lying generates a particular sociocognitive dynamic that has become integral to our experience of literature. Once we become aware of this dynamic, we can approach *Patterns of Childhood* with a different set of questions. We can ask not just what lying does for Nelly and the narrator that she is destined to become, but also what it does for Wolf, the writer, whose task it is to transform her childhood memories into a novel.

To show what lying does for Wolf, this essay will combine cognitive literary theory and archival research.[1] This is an argument with many moving parts, so let me tell you what they are and how they fit together. I will start by discussing the role of complex mental states—which is to say, mental states *about* mental states—in generating literary subjectivity, and the special place of lying in the evolution of this subjectivity. (This is a big topic, but I will keep it brief; for an in-depth treatment, see my book, *The Secret Life of Literature*). I will then narrow this broader conversation about literary subjectivity down to life-writing genres, drawing on converging insights from cognitive science and literary studies about the fundamentally constructed nature of "authentic" memories. I will then offer you a hypothesis about what one may expect to find in an original manuscript of an autobiographical novel if one focuses specifically on the writer's treatment of complex mental states. Finally, I will tell you what I found in the Christa Wolf archive of the Berlin Academy of Arts and discuss the significance of my findings for future archival studies driven by cognitive literary theory.

1. Social Cognition and Literary Subjectivity

Literature both builds on and experiments with our social cognition. One specific way in which it does so is by cultivating situations that require readers to process mental states embedded within other mental states, as, for instance, when we make sense of a character's actions by realizing that she *doesn't want* the other character to *know* what she *is thinking*. The range of social and stylistic nuances used to construct such recursive embedments is practically endless, especially if we remember that the mental states in question are not limited to those that readers attribute to characters but can also arise from the mutual awareness between the implied reader and the implied author.

The strong version of this argument, which I advance elsewhere, is that many literary genres, as we know them today, such as novels, plays, narrative poems, and memoirs concerned with imagination and consciousness, cannot function anymore without constantly eliciting complex mental states in their readers by embedding mental states on

at least the third level (as in, mental state within mental state within yet another mental state). This means, among other things, that when a text is intended to disrupt an easy mentalizing—for instance, when an author writes a novel pointedly devoid of any references to psychological interiority—audiences will still force-read as many complex embedments into the story as the narrative itself, its cultural contexts, and their own past reading histories will allow them. They will do so because intuitively expecting literary subjectivity to be constructed as a series of complex embedments, explicitly spelled out or merely implied by a text, has become our standard experience of literature.[2] This expectation is supported by a variety of cultural institutions that reward their participants for elaborate attributions of mental states (a practice, one should add, which is not at all universal, for different communities around the globe have divergent perspectives on the ethics of discussing openly other people's thoughts and feelings[3]).

Representations of lying have a special relationship with embedment. It seems that in several national literary traditions, one can observe the following pattern. The further back one goes in history, the likelier it is that complex embedments in literature are created by portraying characters who deceive other characters. In contrast, in what we consider more modern literature, complex embedments are created by a much wider variety of social contexts, which include lying but are by no means limited to it.

The reason lying never goes out of fashion, even when a much richer repertoire of embedment strategies becomes available, is that it delivers embedded subjectivity in a particularly effective package. A character who is engaging in deception *wants* the other person to *think* that she (i.e., the deceiver) is *thinking* something other than she is really thinking. Consider, too, the embedment possibilities of self-deception, as well as the complex social emotions attendant upon lying, such as shame and embarrassment (which themselves pack nice wallops of embedment). All said, there is really no reason for an author to let go of such an expedient, time-tested tool as lying, even if the majority of the text's embedments may now come from other contexts. (Of course, writers themselves do not think about it in these terms. In fact, as I show in my study set in an MFA workshop, being consciously aware of embedding complex mental states is detrimental to the writing process.[4])

This, in brief, is the cognitive-theoretical background against which I propose to view Wolf's preoccupation with lying in *Patterns of Childhood*. If literature as we know it today depends on unceasing elicitation of embedded mental states in readers,[5] then a writer working on transforming her childhood memories into a novel is not likely to ignore this effective

representational tool, although she may also combine it with other nuanced and idiosyncratic stylistic strategies for embedment.

This perspective on lying by no means negates or supersedes other perspectives, including those focusing on its thematic roles, such as, for instance, lying as a means of private resistance to a totalitarian regime, or lying as a form of commitment to truth. Here is how Wolf puts it: "The deceit, and the fact that [Nelly] remained conscious of it, as much as her longing for truthfulness, was this perhaps some form of salvation? A vestige of independence, which she was able to resume later?" (224). These are powerful emotional matters, and they are not reducible to embedded mental states that structure them.

But the fact remains that embedded mental states do structure them. For instance, in the sentence that you just read, the narrator *wants* to *believe* that because Nelly *knew* that she was lying, she actually *longed* for truth. So even though we do not consciously register complex embedments of mental states when we read, our experience of a text's meaning and affective impact depends on our subconsciously keeping track of dynamic relationships among its multiple intentionalities.

Let us see now what we can discover about the process of creating these relationships if we compare the original manuscript of *Patterns of Childhood* with its subsequent revisions.

2. Memory as Forgery

Before I tell you what I had expected to find in Wolf's archive and what I actually found, let us acknowledge the fundamentally fraught nature of the project of reconstructing one's past. Nelly may have been "longing for truthfulness" in the middle of lying, and Wolf's narrator, too, may be longing for an authentic glimpse of the child that she once was, but, as she herself puts it, memory is "forgery" (6). That child is gone, and Nelly is a construct designed to solve a *present* problem, that is, to create a plausible bridge between the narrator's past self and the adult that she has become.

Scholars of autobiography have long been aware that the process of remembering the past is shaped both by the emotional needs of the moment in which it is taking place and by the projected visions of the future self. Already in 1985, John Paul Eakin characterized autobiographical writing "as a ceaseless process of identity formation in which new versions of the past evolve to meet the constantly changing requirements of the self in each successive present" (36).

Cognitive neuroscientists share the view of memory as "a fundamentally constructive act" (Schacter 266).[6] Its "core network," they suggest, may have evolved to imagine a future, which is to say, to simulate "future

experiences by recombining elements of past experiences to construct novel event representations" (Schacter 266).[7] Small wonder that, when applied to the past, memory starts churning up "misattributions and false recognitions" (Schacter and Addis S111), "distortions and illusions" (Schacter 266). A quest for an authentic personal past appears delusional once we recall that memory is geared toward "flexibility" and "divergent creative thinking" (Schacter 265) rather than toward recording and storing set impressions.

With this in mind, how should we feel if we discover that some of the memories of lying present in the original manuscript of *Patterns of Childhood* were changed in subsequent revisions? Would this imply to us that the "true" experiences were "doctored" and thus lost their claim to "authenticity"? It seems that, based on what I just said about the fundamentally constructed nature of memory, the answer to this question should be no. After all, the original memories are as much subject to "distortion and illusion" as the revised ones.

And yet, and yet. Knowing that our memories are imperfect self-serving constructs does not stop us from being strongly emotionally attached to them, especially in social contexts in which asserting their truth has immediate practical consequences. Just so, when it comes to autobiography, a genre whose cultural raison d'être is the expectation of access to the authentic self,[8] we may not be in a hurry to abandon that expectation. If anything, we tend toward the opposite. We'd sooner treat a novel as a covert confession of its author's "true" sentiments (e.g., some readers were shocked upon first meeting Vladimir Nabokov's beloved gray-haired wife, Vera: surely, the man who thought up Lolita could be expected to have on his arm a much younger woman!), than be prepared to distrust details of every memory depicted in an autobiography.[9]

The larger question of why it should be so—that is, what combination of our cognitive predispositions, on the one hand, and the value that our culture places on having access to people's "real" mental states, on the other, makes us so keen on holding authors to their "authentic" memories—is worth addressing separately.[10] For the purpose of this essay, let us just acknowledge that there is a genuine and, perhaps unresolvable, tension in our experience of autobiographical fiction, for we are expected to simultaneously appreciate its unrelenting literariness *and* to look for an authentic remembered self. That the genre exists and even thrives is a testament to a ceaselessly exploratory nature of human social cognition: our tendency to carve out bits of habitable ground in unstable conceptual spaces and then keep building there—not toward any kind of discernible goal but just because we can.[11]

So, yes, we may feel somewhat let down if we find that the original memories of Nelly's lying were significantly changed in later versions of the manuscript. After all, the genre of autobiographical novel depends on our belief in "authentic" memories, even if, armed with insights from cognitive science and literary theory, we should know better.[12] So perhaps we would try to mitigate our disappointment by telling ourselves that the revisions indicate better, more accurate memories: for isn't it possible for a writer, as she keeps thinking about her past, to uncover and restore some hitherto-forgotten layers of nuance?

And if those newly uncovered nuances just happen to feature complex embedments of mental states, well, isn't that a pretty accurate reflection of real life? To quote developmental psychologist Michael Tomasello, "human collaborative activity and cooperative communication both rest on . . . recursive intention-reading" (173). For don't we recursively embed our own and other people's mental states as we negotiate our "collaborative" social interactions?

We do, but not at the same rate and with the same elaborate intensity at which literature does it (or prompts us to do). Although throughout the day we may occasionally embed mental states on the third or even fourth level—as in, I *am sure glad* that he *didn't know* how I really *felt* about it, or, I *wonder* if she *expected* that they would *want* to *surprise* her—most of our daily social interactions do not require such complex embedments. We mostly get by with second-level ones, as in, oh, he's heading toward the kitchen, I *wonder* if he *wants* something to drink, or: she is taking a long time getting ready, it's because she doesn't have her watch on, I *guess* she *doesn't know* what time it is. In general, thinking about thinking about thinking (third-level embedment) "occurs in interpersonal cognition in real life less frequently" than, for instance, thinking about thinking (second-level embedment). The former, as social psychologist Patricia Miller and her colleagues put it, "has a lower ecological plausibility" (622).

Literature, by contrast, has learned to make high-level embedments so "ecologically plausible" that we don't even notice them. A 58-word stanza from *Eugene Onegin*, Alexander Pushkin's novel in verse, can make us process fifteen or so third-level embedments in about ten seconds (which is, roughly, the time that it takes to silently read it).[13] Yet, in Russian culture, *Onegin* is not associated with laborious intellectual processing and surreal demands on one's social cognition. On the contrary, it is known for its playfulness, fluidity, and, in some circles, for "simplicity and realism."[14] It is quite ironic that our notion of literary realism, deservedly contested in so many other ways, remains unselfconsciously tethered to massively exaggerated and as such not exactly "realistic" patterns of embedment.

3. A Hypothesis About What One May Find in an Archive

Here, then, is what I hypothesized about what I might discover while working with manuscript versions of Wolf's autobiographical novel. Specifically, when it came to Chapter 10 of *Patterns of Childhood*, that is, the chapter that deals with Nelly's lying, I expected that in later versions Wolf would have added *more* complex embedments to situations featuring lying. I did not expect to see any increase in the *level* of embedment—for instance, from the third to the fifth level—because, as I have argued elsewhere, extremely intricate social nuances can be conveyed on the third and fourth level of embedment.[15]

My expectation that revisions would add more embedments was based, among other things, on my study of conversations that took place in a creative writing seminar, whose participants were expected to read each other's short stories and then revise their work in response to their colleagues' comments. What I observed in that seminar is that, although writers are not familiar with the concept of embedment, when they comment on each other's drafts, their suggestions for improvement tend toward making social situations present in the original more emotionally complex, which, of course, depends on cultivating complex embedments of mental states. Some of those suggestions center on motivations of characters, others involve various states of awareness between the narrator, the reader, and the author. While I do not have enough data to say that all the resulting revisions ended up featuring more complex embedments, enough of them did for me to assume that, in a text as preoccupied with consciousness and imagination as Wolf's *Patterns of Childhood*, new sentences added in the process of revision would contain complex embedments, thus increasing the incidence of such embedments from one draft to another.

I should add that I focused on embedments surrounding lying because of the centrality of this subject to this portion of Wolf's novel. I could have picked some other focal point, e.g., Nelly's relationship with her mother, or the narrator's relationship with her own daughter. Complex embedments tend to cluster around various points of intersubjective tension (although we can also think of it in reverse and say that points of tension are created by clusters of complex embedment). It so happens that lying also has a special status in the context of the history of embedment, for it appears to be the oldest reliable source of complex embedment in many national literary traditions. This adds an extra dimension to my present argument, but it certainly does not mean that researchers interested in doing an archival study along similar lines (i.e., seeing if the pattern of complex embedment changes from one stage of revision to the next) should limit themselves to contexts featuring deception.

4. Compressed Embedments

Chapter 10 of Wolf's novel centers on the fifteen-year-old Nelly's relationship with her history teacher, Dr. Julia Strauch, "the leader of the National Socialist Women's Organization" (219). Nelly adored Julia, and Julia considered Nelly her star pupil. Yet almost every encounter with Julia involved some form of deception. For instance, when Julia wanted to talk to her favorite student "seriously," she used an "encouraging and understanding" tone, "which—as Julia knew only too well, as Nelly knew that Julia knew—had direct access to Nelly's 'innermost being'" (227). Not surprisingly, given Nelly's awareness of Julia's intention to access her "innermost being," when such conversations did take place, both the teacher and the student were entangled in a subtle web of deceit and manipulation.

Thus, when Julia started discussing with Nelly "signs of unruly behavior in her class"—such as "the formation of small cliques," which was particularly unacceptable now that Germany was waging "the decisive battle against her enemies"—Nelly omitted to mention that the "real danger" of "insubordination" involved not the girls who exchanged notes during class, but one "Christa T., the new girl from the Friedeberg area, who didn't show off and "had no need for Julia," and whose friendship Nelly was now eager to win:

> Not a word about that to Julia. Instead, the appearance of being reasonable, as always, and yet, although unintentionally, a touch of distance. Which was enough to make Julia pronounce the sentence Nelly had vainly waited for for so long: Well, you and I know what we mean to each other, don't we? The sentence came too late, no doubt, and had almost lost its magic effect. Nelly wouldn't have admitted it to herself, but she had felt a sneaking suspicion that Julia might be calculating. (228)

If we try to map out recursive embedments present in this passage, we may come up with something along the lines of: Nelly *suspects* that Julia is *aware* that Nelly *does not want* her to *know* what she is *really thinking*. So she *wants* to *remind* Nelly of their mutual *awareness* of their special bond. This of course, makes Nelly *wonder* if they indeed mean to each other as much as Julia *wants* her to *think* they do. As always in such cases, my (crude) map may be different from yours, but what's important is not that our maps should be identical, but that we agree about what we are trying to capture with those maps. And what we are trying to capture is the social complexity of the situation—its meaning, which is irreducibly bound with the interplay of mental states.

Another such passage features Nelly trying to write an essay for Julia's class. As she had discovered, it was easy for her to write school essays "about general familiar everyday notions," such as "a nation hemmed in," or "the Nordic spirit in the poetry of antiquity," but when it came to

such subjects as "the first snow," which "required writing about personal matters," it was "a lot more difficult":

> She remembered clearly: while describing the particular Sunday on which the first snow had fallen that year, she never forgot for a second for whom she was writing. A touch of deceit permeated every line; she had described her family as just a trifle too idyllic and herself as a just a trifle too virtuous: exactly the way she thought Julia wished to see her. (The deceit, and the fact that she remained conscious of it, as much as her longing for truthfulness, was this perhaps some form of salvation? A vestige of independence, which she was able to resume later?) (224)

Note how new complex embedments keep building up from the initial lie. Nelly knows that she is deceiving Julia. This is to say that she *knows* that she is describing herself the way she *thinks* Julia *wants* to see her. Moreover, the narrator *knows* that Nelly *knows* that she *wants* to *deceive* Julia, and she wonders if she owes her present-day subjectivity to Nelly's future memory of her former awareness of her wish to deceive.

What we have in this passage may be best described as *compressed* embedments, which is to say concepts that are themselves already complex embedments and that can be used as building blocks for further embedments.[16] Compressions can be linked with complex social emotions, such as shame and embarrassment; with certain behaviors, such as hypocrisy and deception; and with teleological concepts, such as redemption and salvation. We can think of them as cultural shortcuts that facilitate communication by bundling together recursively embedded mental states associated with difficult yet not uncommon social situations.

Think, for instance, of the recursive scaffolding that makes possible the following embedment: "I am embarrassed about lying to that hypocrite." Embarrassment involves *thinking* about what other people *may be thinking* about one's presumably unappealing *motivations*, while lying presupposes *awareness* of *wanting* someone to *think* something that we *know* is not the case. And when we accuse someone of hypocrisy, what we mean is that the person *wants* others to *think* that he is *committed* to something that, as his behavior reveals, he is not really committed to. "I am embarrassed about lying to that hypocrite" is thus a remarkably efficient way to communicate what is actually an extremely complex social sentiment. The fact that we do not appreciate its underlying complexity shows how transparent (i.e., see-through, invisible) the workings of social cognition can become to us. We see what we want to communicate without noticing what makes the process possible.

As to teleological compressions, they imply some form of an accounting system in the universe, by evoking mental states of spiritual entities, such as God, Providence, karma, etc., that keep track of people's "real" intentions and mete out just desserts based on those intentions. Such

compressions do not have to involve any religious beliefs. For instance, one can be an atheist and still feel that all is not well with the world in which a flagrant wrongdoer is thriving while a good person is miserable. Articulating unspoken assumptions behind such intuitions of cosmic justice can feel uncomfortable, because of how readily they seem to fit familiar religious paradigms. What they may actually represent, however, as cognitive psychologists have shown, is not some inborn religiosity but an extension of our social cognition to events in the world, which can be conscripted by religiosity but doesn't have to. (See Jesse Bering's *The Belief Instinct* for an elegant exploration of cognitive predispositions that may prompt attributions of mental states to events.)

So when Christa Wolf speaks of "salvation," we do not assume that the famous East German writer found the Lord in the 1970s. Instead, what she says makes sense to us because we, too, can readily imagine some overarching account-keeping system. Which is to say that we can imagine (without necessarily articulating it to ourselves this way) some mind capable of seeing the "real" meaning behind people's behavior, a meaning they themselves may not be able to see—and thus be qualified to judge and forgive them. Call it a teleological version of dramatic irony: an *awareness* of someone *knowing* something that the people immediately involved in the situation *do not know* (dramatic irony, incidentally, being yet another cultural shortcut for a complex embedment recurring in certain settings).

To return to Wolf's parenthetical query—"The deceit, and the fact that she remained conscious of it, as much as her longing for truthfulness, was this perhaps some form of salvation? A vestige of independence, which she was able to resume later?"—if we unpack the complex embedments behind its compressions, we end up with a series of painfully unwieldy sentences. Here is one possibility: The narrator imagines a world in which someone (perhaps that narrator?) forgives Nelly for going along with the Nazi regime because that someone knows that Nelly remains aware of presenting herself the way she thinks Julia wants to see her, and that she doesn't really want to present herself this way. In the grand scheme of things envisioned by that omniscient someone, Nelly has to go through the ordeal of being at home in Nazi Germany because it allows her to cultivate a consciousness that would later make her the kind of writer who is capable of thinking independently and thus make that world a better place.

No one would want to read a novel featuring such constructions, but with compressions doing their job in *Patterns of Childhood*, we *are* processing such embedments and enjoying it, too. Chapter 10 of Wolf's novel feeds us complex embedments built out of compressed embedments at a steady rate; we get not just shame, hypocrisy, embarrassment, lying,

manipulation, and teleological thinking, but also various combinations of those: embarrassment about manipulation, shame about hypocrisy, teleological thinking about lying. Difficult, painful, and confusing as these subject matters can be, they are structurally helpful for a writer working on transforming her memories into literature.

Now, finally, let us see when and how they make their first appearance in the text.

5. What One Finds in an Archive

The Christa Wolf archive in the Akademie der Künste in Berlin contains close to 175,000 sheets of her manuscripts, diaries, and correspondence. In the case of *Kindheitsmuster,* the archive has two typewritten manuscripts, an early one and one mostly identical to what went to print, as well as handwritten notes. I thus worked with two typewritten versions of Chapter 10—File 562 (the early version, 1973-1974) and File 580 (the later version, 1975), focusing, in particular, on the changes in File 580, in the passages that deal with Nelly's relationship with Julia. What I found was that, indeed, these changes all featured new complex embedments, subtly transforming the meaning of the interactions between the teacher and the student.

It is worth noting that when we first meet Julia, Wolf's narrator is uncharacteristically sanguine about her powers of remembering: "As for Juliane Strauch . . . your memory couldn't be more exact. Her face, her figure, her walk and behavior have been preserved within you for twenty-nine years" (218). How should one reconcile this assertion of accurate recall with the particular type of revisions that one then finds in the manuscript—revisions that add complex embedments and, by doing so, significantly modify the emotional dynamic of those "exact" memories?

Perhaps no such reconciliation is possible. Our search for an "authentic" autobiographical self is perennially at odds with our appreciation of literary subjectivity generated specifically by embedments. The tension between the two is always there, but archival research of the kind that I describe here, which is to say a study of manuscripts prompted by insights from cognitive literary theory, offers new ways of bringing it to the surface.

Case Study I: Nelly Knew that Julia Knew

We start by revisiting the "serious talk" that Julia has with her favorite student about the "signs of unruly behavior in her class":

> She didn't reprimand, but remained encouraging and understanding. Exactly the tone which—as Julia knew only too well, as Nelly knew that Julia knew—had direct access to Nelly's 'innermost being'" (*Patterns of Childhood* 227).

Genau der Ton, der—was Julia natürlich wußte, und Nelly wußte, daß sie es wußte—einen directen Zugang zu Nellys „Innerem" hatte. (Kindheitsmuster 291)

In the original version of this last sentence, Julia already uses the tone which, she knows, will give her direct access to Nelly's innermost being, but the phrase "Nelly knew that Julia knew" is not there: "*Derselbe Ton, für den es, wie Julia natürlich wußte, einen direkten Zugang in N.s 'Inneres' gab, in dem Julia gelesen haben muß wie in einem aufgeschlagenen Buch*" (File 562, p. 338; Figure 1). The original version also has the description of Julia reading Nelly's innermost being "like an open book" ("*wie in einem aufgeschlagenen Buch*"), which the later version drops.
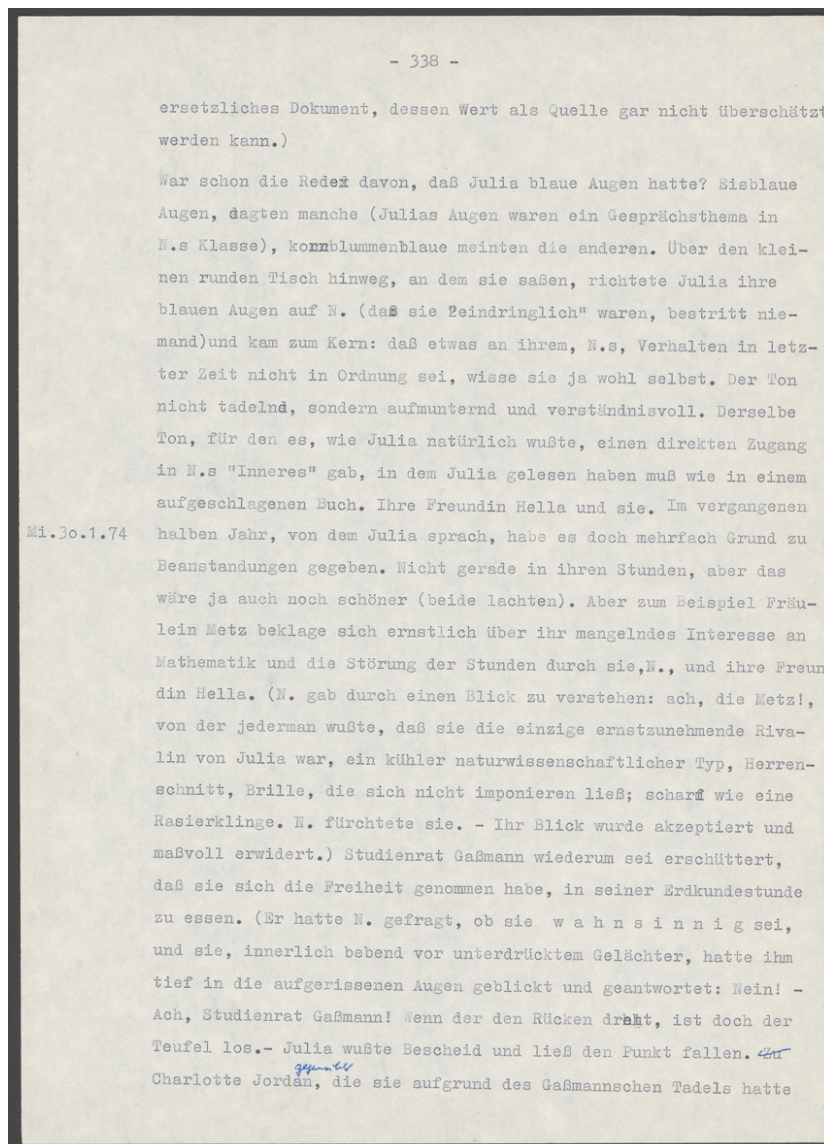

Figure 1. File 562, p. 338 ("Wolf-Christa 562")

One effect of adding "Nelly knew that Julia knew" ("*Nelly wußte, daß sie es wußte*") is that it makes Nelly aware of Julia's intentions. This, of course, also renders Nelly less of "an open book": if she knows that Julia knows which tone to use to best manipulate her, then Nelly is not *completely* readable to Julia. And, crucially, it is this heightened new awareness of her own and of other people's mental states that connects the child Nelly with the writer that she is destined to grow into. Forging that connection is an important project for Wolf's narrator who (to quote John Paul Eakin again) is hard at work creating a "new version of the past" to meet the evolving requirements of the present self (36).

Case Study II: She Had Felt a Sneaking Suspicion that Julia Might be Calculating

We see similar a dynamic at work in the episode describing Julia's reminder to Nelly that they "know what [they] mean to each other," when Julia intuits that Nelly, secretly fascinated by the independent spirit of "the new girl from the Friedeberg area," begins to draw away from her.

The original phrasing, contained in File 562 (Figure 2) is as follows:

> Not a word about that to Julia. Instead, the appearance of being reason-able, as always, accepting her farewell greeting: Well girl, you'll do it.

> *Darüber kein Wort zu Julia. Stattdessen verständig sich zeigen wie immer und ihren Abschiedsgruß entgegennehmen: Nun Mädel, du machst das schon.*
> (File 562, p. 339)

The difference between this and what one then encounters in File 580 is striking, especially if one keeps in mind the narrator's above-mentioned assertion that, when it comes to Julia's behavior, her "memory couldn't be more exact" (218). Julia's short farewell, "Well, girl, you'll do it" (which I interpret as, "you'll do the right thing," or, "you're already doing the right thing"), is transformed into an elaborate description of mutual manipulation:

> Not a word about that to Julia. Instead, the appearance of being reason-able, as always, and yet, although unintentionally, a touch of distance. Which was enough to make Julia pronounce the sentence Nelly had vainly waited for for so long: Well, you and I know what we mean to each other, don't we? The sentence came too late, no doubt, and had almost lost its magic effect. Nelly wouldn't have admitted it to herself, but she had felt a sneaking suspicion that Julia might be calculating. (*Patterns of Childhood*, 227-228)

> *Darüber zu Julia kein Wort. Stattdessen verständig sich zeigen wie immer, und doch, wenn auch absichtslos, mit einem Anflug von Distanz. Der reichte aus, um aus Julia Satz herauszuholen, auf den Nelly so lange vergeblich hatte warten müssen: Wir beide wissen ja, was wir aneinander haben, nicht wahr?*

Der Satz kam zu spät, zweifellos, und er hatte seine Wunderwirkung beinahe schon eingebüßt. Nelly hätte es sich nicht zugegeben, aber es ~~war~~ kam der Verdacht in ihr auf~~gekommen~~, Julia könnte berechnend sein. (File 580, p. 299).

I have already discussed (in the preceding section) some of complex embedments contained in this passage. Here I only want to point out that one effect of those new embedments (which, incidentally, all explore different shades of deception) is to deepen Nelly's intersubjective awareness and, by doing so, establish further continuity between her and the future narrator of *Patterns of Childhood.*
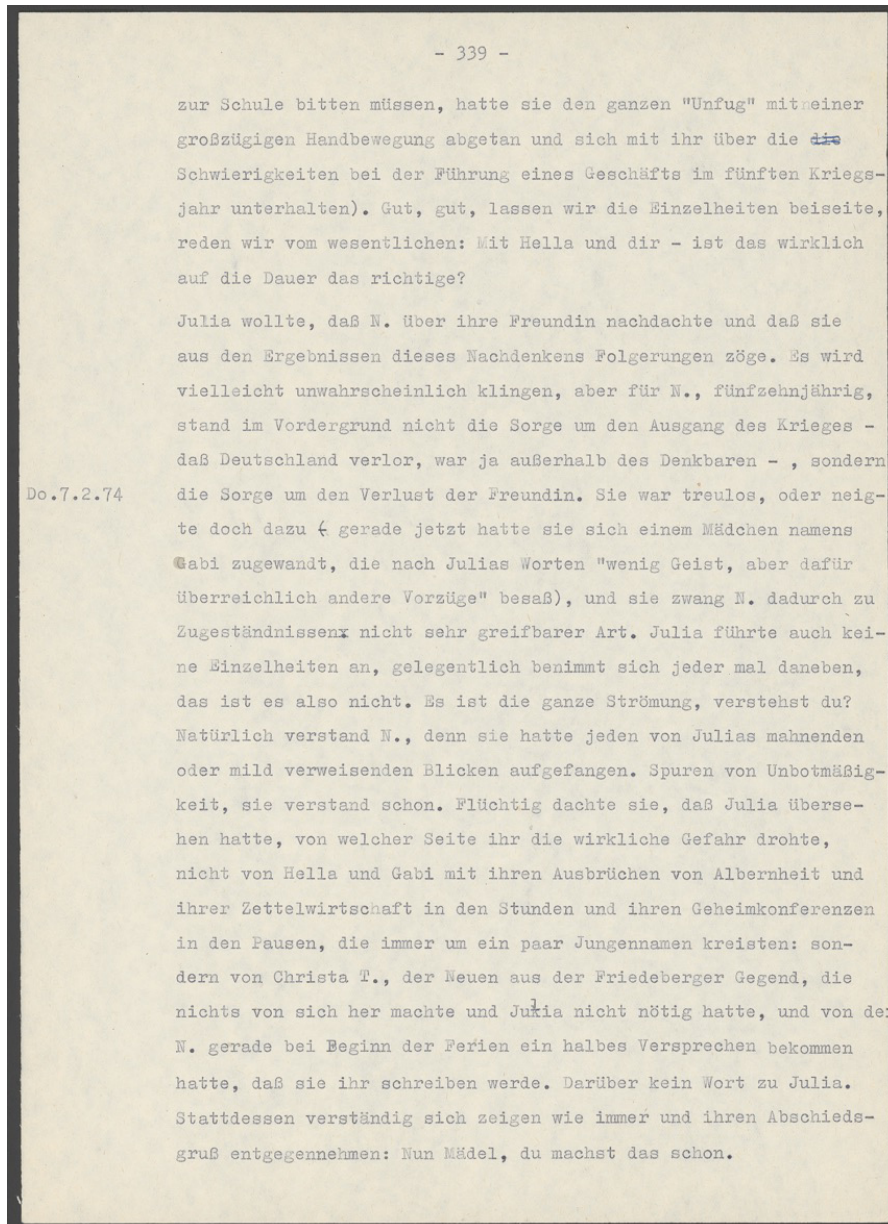

Figure 2. File 562, p. 339 ("Wolf-Christa 562")

Case Study III: The Deceit, and the Fact that She Remained Conscious of It . . .

Here is another example. Recall Nelly's difficulties with writing personal essays: her awareness of portraying herself "exactly the way she thought Julia wished to see her," and hence of the "deceit" inherent in her description of "her family as just a trifle too idyllic and herself as a just a trifle too virtuous" (224). This passage was heavily revised in its transition from File 562 to print (see Figures 3 and 4), but to make our discussion of it manageable, I will focus here only on two revisions.

Here is how it read originally:

> . . . exactly the way she thought Julia wished to see her. She also knew that in order to deceive Julia - or to win her over, which seemed to be the same thing - she would have to refrain from all clumsy maneuvers and use the finest and most subtle arts of disguise, of the kind that, even if they are noticed, must still flatter the one to whom they are directed. A web, with which N. thought to ensnare Julia, looks, gestures, words, lines, which lay a hair's breadth beside the sincere feelings, but never quite coincided with them. [25]

> *Genauso, wie sie glaubte, daß Julia sie zu sehen wünschte. Sie wußte auch, daß sie, um Julia zu täuschen, - oder zu gewinnen, was dasselbe zu sein schien – sich aller plumpen Manöver enthalten und feine und feinste Verstellungskünste anwenden mußste, von der Art, die, selbst wenn sie bemerkt werden, dem, an den sie sich richten, noch schmeicheln müssen. Ein Gespinst, mit dem N. Julia zu umgarnen dachte, Blicke, Gesten, Worte, Zeilen, die haarscharf neben den aufrichtigen Empfindungen lagen, doch niemals ganz mit ihnen zusammenfielen.* (File 562, p. 335)

And here is how it reads in the published version:

> . . . exactly the way she thought Julia wished to see her. (The deceit, and the fact that she remained conscious of it, as much as her longing for truthfulness, was this perhaps some form of salvation? A vestige of independence, which she was able to resume later?)
>
> In order to win Julia over—or to deceive her, which seemed to amount to the same thing—she had to refrain from blunt maneuvers and ensnare the demanding teacher, who was not easily flattered, in a wave of the subtlest weave; looks, gestures, words, lines that lay within a hair's breadth of her true emotions, without ever fully blending with them. (*Patterns of Childhood*, 224-225)

> *Genauso, wie sie glaubte, daß Julia sie zu sehen wünschte. (Die Heuchelei und daß sie ihr schwach bewußt blieb, ebenso wie die Sehnsucht nach Aufrichtigkeit: Vielleicht war das eine Art von Rettung? Ein Rest von Eigenleben, an der sie später anknüpfen konnte?)*
>
> *Um Julia zu gewinnen—oder zu täuschen, das schien dasselbe zu sein hatte sie sich aller plumpen Manöver zu enthalten und die anspruchsvolle Lehrerin, der nicht leicht zu schmeicheln war, mit einem Gespinst feinster Art zu umgarnen: Blicke, Gesten, Worte, Zeilen, die haarscharf neben den aufrichtigen Empfindungen lagen, doch niemals ganz mit ihnen zusammen-fielen.* (*Kindheitsmuster*, 288)

Let us take a closer look at the important parenthetical insertion, "The deceit, and the fact that she remained conscious of it, as much as her longing for truthfulness, was this perhaps some form of salvation? A vestige of independence, which she was able to resume later?" This insertion adds a new set of embedments (including some compressions)—to a reflection that was complex to begin with. This is what we had here initially: Nelly *knows* that she describes herself in her essays the way Julia *wants* to *think* of her. What we have now is: the narrator *hopes* that Nelly's *awareness* of
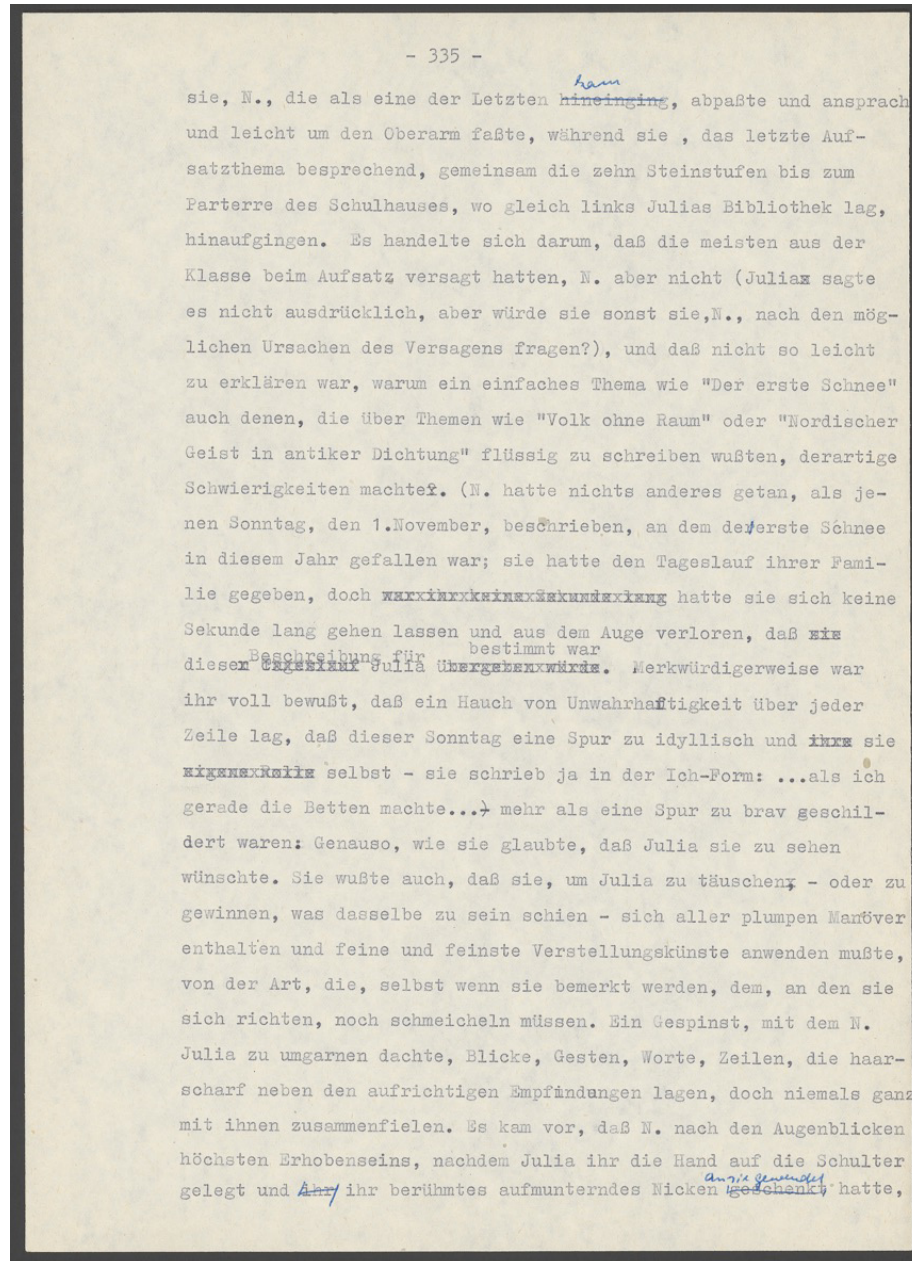

Figure 3. File 562, p. 335 ("Wolf-Christa 562.")

describing herself the way Julia *wants* to *think* of her transcends the immediate situation, because this awareness connects the child to the writer and independent thinker that she was destined to become. If, driven by a perverse critical insistence on looking behind the curtain, we unpack the compression implied by the word "salvation" and ask *whose* perspective this word implies, we get something painfully unwieldy, along the lines of: the narrator *hopes* that something in the universe is *keeping tabs* on Nelly's *awareness* that she describes herself the way Julia *wants* to *think* of her.
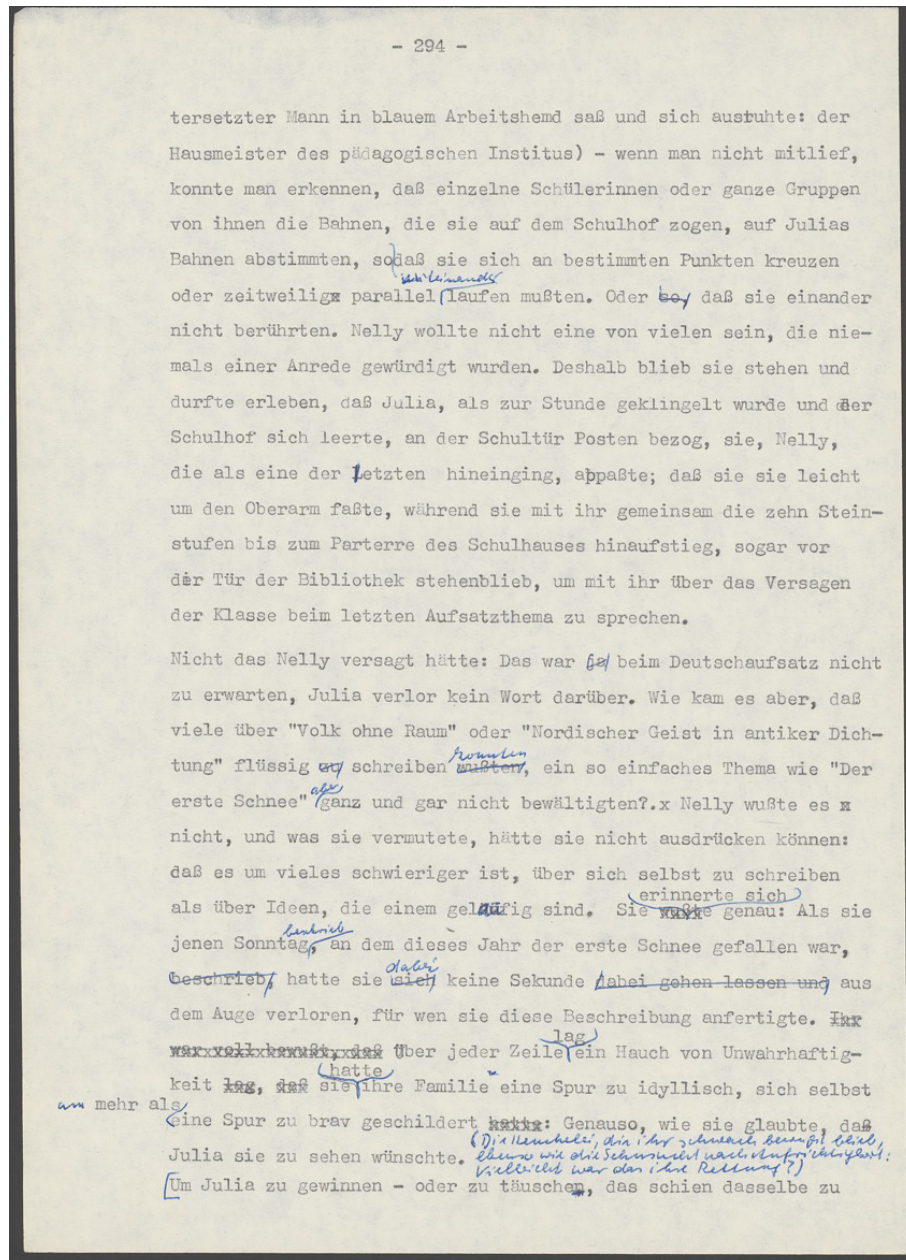

Figure 4. File 580, p. 294 ("Wolf-Christa 580")

To emphasize: we do not consciously spell out these embedments when we read. Yet to make sense of the extremely complex social situation conjured up by the text, we have to process them, just as we have to process various embedments structuring our real-life social interactions. We may not quite grasp them, and we may misinterpret them, and we may experience them somewhat differently depending on our respective personal histories, but we can't avoid relying on recursive mindreading (which is transparent to our social species and hence requires a special cognitive and critical effort to notice it).

What is interesting about this particular manuscript revision is that, in its entirety, it represents not just an afterthought, but an after-afterthought, as it were. The sentence, "The deceit, and the fact that she remained conscious of it, as much as her longing for truthfulness, was this perhaps some form of salvation?" is not there in File 562, but *it is also not there* in the typed File 580! Instead, it is *handwritten* into File 580 (see Figure 4), which means that it was inserted at a later stage. Even more interesting, the sentence, "A vestige of independence, which she was able to resume later?" *is not there in File 580 even in the handwritten form*, which means that it was inserted at a yet later stage. What this all adds up to is that Wolf kept changing this part—returning to it on at least three occasions—and on each occasion, her revisions introduced new complex embedments. What started as a relatively straightforward description of deceit, has turned into a *reflection* on deceit, building on at least two compressed embedments: deceit itself and a teleological concept of salvation.

6. Notes on Future Research

Such, then, is at least one aspect of the process of transforming memories into literature, which is to say, transforming one kind of imagined reality into another. Whatever the "original" memories may be, we can assume that in the course of becoming literature, they accumulate more complex embedments. Most of that change happens off-page, but some of it is part of a written record, if we are lucky enough to have access to original manuscripts. What is exciting about cognitive literary theory is that it thus allows us to enter an archive with a new set of questions about creative writing and have those questions answered.

With time, I hope that we will see more archival queries inspired by insights from cognitive science. For instance, I look forward to studies inspired by the respective research of cognitive literary scholars Gabrielle Starr and Laura Otis, which would trace changes in patterns of multisensory imagery from one manuscript stage to another. Do writers add more sensory modalities to some situations, for instance, introducing taste and touch where initially there were only sight, sound, and smell? Or do they

integrate those modalities in a new way? Or do they, perhaps, remove some senses from the mix, strategically impoverishing some characters' embodied experience of their worlds? I ask these hypothetical questions to show you that a cognitive-archival inquiry does not have to be limited to the exploration of embedded subjectivity, even though this is the subject that I am personally most excited about.

With an eye on future studies, here are two issues that we may want to clarify. First, I have shown that, in the process of revision, Wolf added more complex embedments to her text, sometimes increasing the characters' capacity for mutual awareness, sometimes superimposing a perspective of the narrator onto that of a child character, sometimes using compressed embedments, such as lying, as building blocks for further embedments. But, one may ask, do manuscript changes always result in *adding* complex mental states? Isn't it also the case that some revisions may *remove* complex embedments?[26] The answer to this question is yes. I did come across several paragraphs in File 562 that did not make it to File 580, which means that all complex embedments contained in those paragraphs were gone.

Here is something to keep in mind when we think of such cases. Any revision that gets rid of sentences and paragraphs is likely to result in the loss of some complex embedments, for the simple reason that, in works of literature, sentences and paragraphs are saturated with embedded mental states. To put it differently, when we get rid of a sentence, we are likely to lose some grammatical constructions, because grammar is integral to sentence structure. So (or nearly so) it is with embedded mental states in literature. They are not just extraneous embellishments which can be extracted from the fabric of writing without changing its meaning. As the underlying grammar of literary subjectivity, they live and die with the text.

That said, some cuts *add* complex embedments, by prompting readers to think harder about characters' motivation in order to make sense of what is going on. For instance, scaling down or completely eliminating *explicit* references to thoughts and feelings forces readers to start figuring out nuanced complex embedments *implied* by social situations in which characters find themselves. In general, it is important to remember that cuts do not lead to *simplifying* social situations. Writers may eliminate details that make those situations easy to comprehend, all the while retaining or even increasing their underlying sociocognitive complexity. What this means is that writers known for their ruthless cutting practices may be particularly inviting subjects for cognitive-archival investigation into the dynamics of complex embedment.

Finally, some paring down of embedment may also be strategic. I do not have any examples of this from *Patterns of Childhood*, but I am aware of eighteenth-century British and Chinese novels, in which certain characters are portrayed as less capable of embedding complex mental states than other characters.[27] It is not inconceivable, then, that an author who wants to portray a character of (for instance) a particular class, gender, race, or age as less emotionally complex and socially perceptive than others, would intuitively scale down that character's ability to entertain complex embedments, and that we, working in the archive, would be able to trace that kind of change from one manuscript stage to another.

On the whole, I do expect to see that manuscript revisions, especially when it comes to autobiographical novels or memoirs concerned with imagination and consciousness, would result in adding more complex embedments to the text. But these additions would always be integrated with thematic and stylistic imperatives, for instance, with a narrator's attempt to bridge the gap between the child that she used to be and the adult she has become. Increasing instances of complex embedment is not a goal that a writer consciously pursues,[28] even if she does end up adding more embedments and developing innovative ways of constructing them.

Another issue to consider is the relationship between complex embedment and reflexivity, especially given the genre I am working with here and the particular text. Wolf's autobiographical novel is a meditation on memory, writing, *and* self-awareness. It is so intensely metacognitive that one begins to wonder if literary reflexivity and complex embedment are essentially interchangeable terms. And if they are interchangeable, would not literary critics save themselves the trouble of learning the new concept (i.e., embedment) by simply falling back on the familiar one? After all, reflexivity as well as self-reflexivity have been associated with literature since antiquity,[29] and the centrality of self-reflexivity to the "rise" of the novel has been extensively discussed by literary scholars.[30]

Reflexivity and self-reflexivity are interesting terms when it comes to social cognition because they are, themselves, compressed embedments. This means that, along with irony and dramatic irony, they function as efficient cultural shortcuts when critics want to discuss certain recurrent patterns in literature without droning on, "the author *wants* us to *know* that she is *aware* of her character's *wanting* to *know* . . .", etc. But, useful as these terms are in thus reducing the cognitive load on critics and their audiences, they are not interchangeable with complex embedment because they represent its subset, albeit a significant one. This is to say that not every novel is (self)reflexive, but we would be hard put to find a novel

that does not extensively rely on complex embedment of mental states (even if some texts make a point of eliminating all explicit references to mentalizing).

Along the same lines, I should say that even the autobiographical novel, gravitating as it may be, on the whole, toward greater (self)reflexivity would not rely exclusively on it to cultivate its complex embedments. A researcher working in an archive thus should not assume that instances of complex embedment will cluster predominantly around moments of metacognitive reflection. Reflexivity is comparable in this respect to lying. Lying has been generating complex embedment in literature since at least *The Epic of Gilgamesh*, and it is still going strong, yet we would not expect it to be the *only* source of embedment in a text (including one as preoccupied with pretense and manipulation as Wolf's chapter on Nelly's relationship with Julia Strauch). Reflexivity and self-reflexivity, lying, shame, hypocrisy, and teleological reasoning are all powerful tools for generating social situations that mimic our real-life mindreading challenges but also take us on wild cognitive rides that go beyond those daily challenges. To follow the construction of those rides, to see how they gain in speed and intensity without appearing unwieldy[31] or lopsided[32] is an exciting privilege newly afforded by cognitive-archival research.

University of Kentucky

Acknowledgment

Notes

1. These two approaches have been successfully combined before; see, for instance, Stephenson, *Performance*; Morgan, "Varieties of Freedom"; and Morgan and Burns, "How to Write."
2. See Zunshine, *The Secret Life*, p 221.
3. For a discussion of cultures of opacity, see Zunshine, *The Secret Life*, pp. 113-125.
4. See Zunshine, *The Secret Life*, section 1.12: "What Do Writers Actually Say When They Talk about the Secret Life of Literature?" pp. 39-43.
5. See Zunshine, *The Secret Life*, chapter 1.
6. As cognitive literary critic Alan Richardson puts it, "The close links postulated between remembering and imagining may also help account for the well known fragility and error-proneness of episodic (autobiographical and richly detailed) memory, which may be adaptively designed less for accurately recollecting the past than for creatively modeling the future" (278).
7. See, however, Richardson's warning that "adaptive arguments demand special critical scrutiny," as well as his demonstration how works of fiction can "suggest that the 'adaptive significance' of a flexible, future-oriented, and fragile memory system can readily be overstated (286).
8. Although, as Naomi Rokotnitz reminds me, actual specimens of the genre may end up offering, instead, justification or obfuscation.
9. See Elfenbein for a discussion of readers' belief "that all narrative literature is life writing in disguise" (this volume).

10. For a discussion of a culture-specific belief that there are such things as "true" mental states and that they are knowable, see Zunshine, *The Secret Life*, pp. 138-139.

11. For a related discussion, see Spolsky, *Gaps in Nature*, chapter 2, and *The Contracts of Fiction*, pp. 73 ff.

12. For a useful rethinking of our traditional notion of "authenticity," from a cognitive perspective, see Rokotnitz.

13. For a discussion, see Zunshine, *The Secret Life*, pp. 142-143. Note that the stanza in question, "*Как рано мог он лицемерить*," can be longer in an English translation. For instance, Vladimir Nabokov's version, "How early he was able to dissemble," is 74 words long.

14. As Karen Petrone observes, starting from the 1930s, the official Communist Party line was that the "traits of simplicity and realism linked Pushkin to the prevailing literary style of the time, socialist realism" (117).

15. See Zunshine, *The Secret Life of Literature*, p. 36.

16. Compare to a dynamic of compression as enabling "human-scale representations of the otherwise diffuse patterns" (21), as discussed by Mark Turner.

17. Translations are mine, although I try to keep them as close as possible to the ones by Molinaro and Rappolt.

18. I am grateful to Stefka Eriksen and Naomi Rokotnitz for bringing this issue to my attention.

19. See Zunshine, "Mindreading and Social Status" and "Bakhtin, Theory of Mind, and Pedagogy," and "From the Social to the Literary."

20. In fact, if it is a goal, then it may be a recipe for artistic failure. For a discussion, see Zunshine, *The Secret Life*, 39-40.

21. See Lobate and Dominguez, as well as Huber, Middeke, and Zapf.

22. For a review, see Christoph Henke.

23. For an example of unwieldy embedment, see any of my mindreading "maps," e.g., "Nelly *suspects* that Julia is *aware* that Nelly *does not want* her to *know* what she is *really thinking*."

24. By lopsided, I mean leaning too heavily on one particular device.

Works Cited

Bering, Jesse. *The Belief Instinct: The Psychology of Souls, Destiny, and the Meaning of Life*. Norton, 2011.

Eakin, Paul John. *Fictions in Autobiography: Studies in the Art of Self-Invention*. Princeton University Press, 1985.

Elfenbein, Andrew. "Life Writing, Race, and Reader Engagement." *SubStance*, vol. 51, no. 3, 2022. 35-53.

Henke, Christoph. "Self-Reflexivity and Common Sense in A Tale of a Tub and Tristram Shandy: Eighteenth-Century Satire and the Novel." *Self-Reflexivity in Literature*. Edited by Werner Huber, Martin Middeke, and Hubert Zapf, Königshausen & Neumann, 2005, pp. 13-38.

Huber, Werner, Martin Middeke, and Hubert Zapf, editors. *Self-Reflexivity in Literature*. ☐ Königshausen & Neumann, 2005.

Lobate, Jesus Hernandez, and Oscar Prieto Dominguez, editors. *Literature Squared: Self-Reflexivity in Late Antique Literature* (*Studi E Testi Tardoantichi*: Profane and Christian Culture in Late Antiquity, Band 18). Brepols Publishers, 2020.

Miller, Patricia H., Frank S. Kessel, and John H. Flavell. "Thinking about People Thinking about People Thinking about...: A Study of Social Cognitive Development Child Development," *Child Development*, vol. 41, no. 3, Sept. 1970, pp. 613-623. doi: 10.2307/1127211.

Morgan, Ben. "Varieties of Freedom: Eckhart, Seuse, ›Sister Catherine‹ and the Embodied Cognition Research Programme." *Meister-Eckhart-Jahrbuch*, vol. 12, pp. 2018, pp. 275-91.

Morgan, Ben, and Niamh Burns. "How to Write a Phenomenology of Religious Life: Charles Taylor, William James, Martin Heidegger, Gerda Walther and a Practical Example from the Manuscript Transmission of the >Sister Catherine< Treatise." *Religiöse Selbstbestimmung*, 2020, pp. 281-315.

Otis, Laura. *The Neuroscience of Craft*, under review.

----. "The Role of Multimodal Imagery in Life Writing." *SubStance*, vol. 51, no. 3, 2022, pp. 115-131.

Petrone, Karen. *Life Has Become More Joyous, Comrades: Celebrations in the Time of Stalin.* Indiana University Press, 2000.

Richardson, Alan. "Memory and Imagination in Romantic Fiction." *The Memory Process: Neuroscientific and Humanistic Perspectives*, edited by Suzanne Nalbantian, Paul M. Matthews, and James L. McClelland, MIT Press, 2010, pp. 277-296.

Rokotnitz, Naomi. "'Colluding in the conspiracy of their fiction:' Role-Play, the Construction Theory of Emotions, and Relational Authenticity." *Transatlantic Cognitive (Neuro)Cultures*, edited by Meindert Peters and Shannon McBriar, special issue of *Symbiosis* – A *Journal of Transatlantic Literary and Cultural Relations,* vol. 25, no. 2, 2021, pp. 183-211.

Schacter, Daniel L. "Implicit Memory, Constructive Memory, and Imagining the Future: A Career Perspective." *Perspectives on Psychological Science*, vol. 14, no. 2, 2019, pp. 256-272.

Schacter, Daniel L., and Donna Rose Addis. "Remembering the Past to Imagine the Future: A Cognitive Neuroscience Perspective." *Military Psychology*, vol. 21, Supp.1, 2009, pp. 1532-7876, S111.

Spolsky, Ellen. *Contracts of Fiction: Cognition, Culture, Community*. Oxford University Press, 2015.

---. *Gaps in Nature: Literary Interpretation and the Modular Mind*. SUNY Press, 1993.

Stevenson, Jill. *Performance, Cognitive Theory, and Devotional Culture: Sensual Piety in Late Medieval York*. Palgrave, 2010.

Starr, G. Gabrielle. "Multisensory Imagery." *Introduction to Cognitive Cultural Studies*, edited by Lisa Zunshine. Johns Hopkins University Press, 2010, pp. 275-289.

Tomasello, Michael. *Origins of Human Communication*. The MIT Press, 2010.

Wolf, Christa. *Kindheitsmuster: Roman*. 1976. Deutscher Taschenbuch Verlag, 1995.

---. *Patterns of Childhood*, translated by Ursule Molinaro and Hedwig Rappolt, The Noonday Press, 1980.

"Wolf-Christa 562." *Christa-Wolf-Archiv*, 01.01. Prosa, Kindheitsmuster [4. Fassung] 10. Kapitel. Bl. 311-346. 29. Dez. 1973 - 14. Feb. 1974. Academie der Künste, Berlin. Robert-Koch-Platz. https://archiv.adk.de/objekt/2429946

"Wolf-Christa 580." *Christa-Wolf-Archiv*, 01.01. Prosa, Kindheitsmuster [5. Fassung, entspricht weitgehend der Druckfassung] 10. Kapitel. Bl. 276-304. 9.-15. Sept. 1975. Academie der Künste, Berlin. Robert-Koch-Platz. https://archiv.adk.de/objekt/2430290

Turner, Mark. "Compression and Representation." *Language and Literature*, vol. 15, no.1, 2006, pp. 17–27. doi: 10.1177/0963947006060550.

Zunshine, Lisa. "Bakhtin, Theory of Mind, and Pedagogy: Cognitive Construction of Social Class." *Eighteenth-Century Fiction*, vol. 30, no. 1, 2017, pp. 109-26.

---. "From the Social to the Literary: Approaching Cao Xueqin's *The Story of the Stone* (Honglou meng 紅　夢) from a Cognitive Perspective." *The Oxford Handbook of Cognitive Literary Studies,* Oxford University Press, 2015, pp. 176-96.

---. "Mindreading and Social Status." *Further Reading,* edited by Matthew Rubery and Leah Price, Oxford University Press, 2020, pp. 257-270.

---. *The Secret Life of Literature*. The MIT Press, 2022.