

Philosophers as Unreliable Narrators

Audrey Thompson
University of Utah

The notorious distance between many academic writers' on-the-page personae and their in-person personalities is widely thought to be a good thing. If, in person, many of us are quarrelsome, neurotic, egotistical, oblivious, thin-skinned, competitive, self-aggrandizing, and self-indulgent, yet on the page we appear gentle, thoughtful, courageous, well-informed, and touchingly vulnerable, then at least the research has not been contaminated by the personality. In daily life, we may or may not want to know the person behind the prose as a friend, colleague, boss, neighbor, lover, or presiding judge, but in our studies we expect to hear from a distinctive writing persona, someone calmer, more reliable, more open-minded than most of us ordinarily are. As philosophers, we are supposed to aspire to a higher writing self.

I want to argue for taking up not a higher but rather a more modest, troubled self: the writing persona of the unreliable narrator. In fiction, the unreliable narrator is distinct from the omniscient or objective narrator in being a character in her own right, someone with a limited and flawed perspective on the events narrated. Sooner or later — whether through a distinctive narrative voice, contradictions and tensions in the story, accidental self-betrayals, or direct confession — the unreliable narrator reveals that the story being told is not to be taken on trust. In scholarly work, there is a possible parallel to the unreliable narrator: a writing persona that performatively draws attention to the troubling terms of the writer's trustworthiness.

PERFORMING OBJECTIVITY

Had I begun this essay with a more overtly Austenesque opening sentence — say, “It is a truth universally acknowledged that a philosopher who is cranky and opinionated needs to work at sounding more omniscient” — I might have been able to trade on Jane Austen's irony to underscore the discrepancy between the universalism being proclaimed and the reliability of the claim. When Austen says, “It is a truth universally acknowledged, that a single man in possession of a good fortune, must be in want of a wife,” novice readers may see the opening line as a claim to be taken at face value.¹ For experienced readers familiar with the conventions of literary voice, however, the statement represents a point of view that signals its limitations even as it proclaims its universality.

In *Pride and Prejudice*, the narrative voice is ironic but not otherwise unreliable, and the narrator is outside the action. An *unreliable narrator* is usually a character *in* the story who also tells the story, shaping it through egocentrism or youth and naïveté, madness, lack of information, loyalty, infatuation, or misplaced social values. Whereas the author is assumed to know the whole of the story, the unreliable narrator gives us a skewed version. Wayne Booth defines the unreliable narrator as “a profoundly confused, basically self-deceived, or even wrong-headed or vicious reflector.”² His is a more negative reading of the unreliable narrator than I have in mind. Some unreliable narrators are indeed profoundly self-deceptive, self-

indulgent, vicious, or shallow. Some turn out to be murderers; others are Difficult with a capital D, although not actively dangerous. Presumably, philosophers might be among these, but I want to focus on philosophical parallels to more humble, more or less engaging unreliable narrators. Such narrators include Nick Carraway in *The Great Gatsby*, Huck Finn, and the unnamed speaker in Toni Morrison's short story, "Recitatif." Narrators such as these may misinterpret the events they describe, fail to grasp the importance of salient details, or impose a problematic pattern on what they observe. The interpretation they impose on events is not an interpretation to be taken on faith.

In employing one or more unreliable narrators, authors of fiction call attention to the stories beyond the stories that the narrators offer us. Whether early or late, at some point readers begin to realize that the narrator's claims on our trust and sympathy are at best partial and that her assumptions and interpretations are shaping what we know of the narrative; we may be led to wonder how another narrator might have told the tale. Emphasizing the role played by interpretation, the unreliable narrator engages readers as actively knowing and non-knowing participants in meaning-making. Insofar as readers are led to accept problematic judgments on the part of the narrator, for example, we are brought to a realization of our own role in shaping the story we thought we were merely following.³

The unreliable narrator thus presents a performative challenge to passive participation in reading conventions. In scholarly work, a parallel move can be found in feminist sociologist Liz Stanley's essay, "A Referral Was Made: Behind the Scenes during the Creation of a Social Services Department 'Elderly' Statistic," which presents us with a conventional academic case study written in the third-person impersonal, followed by a postscript written in the first-person feminist. In the appended section, Stanley reveals material that she had suppressed in the earlier part of the chapter. Specifically, she tells us that the case study discussed in the conventional account refers to herself and her family. Because the impersonal academic narrative (originally a free-standing paper given at a conference) was written for a nonfeminist sociological audience, Stanley says, she was both "unwilling to discuss" her intimate involvement in the case study and persuaded that she could not take up certain feminist themes and issues as she would want to do, lacking a "feminist form."⁴

Stanley's interruption of objectivity is meant to highlight the complex realities that may be suppressed in the name of objectivity. For the right kind of audience, with the right kinds of understandings and commitments, she says, we can risk rhetorical forms that express and examine the realms of the subjective that have been artificially left out of the "official" documents. Her subjective truths are not meant to be final truths: in giving us a more intimate reading of her family case study, she explicitly acknowledges the dimensions of her parents' lives that she never knew, misread, or was never meant to know. All of this seems to me powerful and provocative. Yet we are still left with the terms set by the dichotomous framing of objectivity. Not only is the value of a reclaimed subjectivity set *against* that of objectivity, as the private reality behind the public facade, but it clings to many of

the values traditionally claimed for objectivity, including the assurance readers are offered that we are reliable judges of what we are reading.⁵

What is distinctive about her essay is less what Stanley asserts in the postscript than what she *performs* by adding the postscript for feminist readers. Having set up feminist readers to read as nonfeminists, she then interrupts our reading-social-science-as-usual. Undermining our borrowed scientific assumptions, she reminds us how easily we are seduced by the rhetorical conventions of our disciplines. When addressed *as if* we were an audience with certain kinds of values and assumptions and investments, we often respond as if we were indeed that kind of audience.

Taking my lead from this performative move, I want to suggest that those of us who struggle with the seductions of an authoritative discourse might borrow the literary convention of the unreliable narrator to trouble some of our professional faith in our ability to pin down the truth, some of the authority accorded to both “objective” and “subjective” academic voices (dichotomously constructed), and our assurance that we as readers know when and how to trust particular authorial voices. By employing an unreliable narrator, we invite readers to engage both with our own intellectual and political investments and with those of our discipline. Note, however, that in novels, short stories, poetry, and film, an unreliable narrator is used by an author fully aware of all that the narrator character refuses or fails to see, whereas in philosophy the fully knowing author herself is to be called into question. Any appeal to an unreliable narrator thus creates a paradox for progressive philosophers. To make use of such a device, we need to know more than we do about the limitations of our understanding.

In this discussion, I am primarily addressing antiracist white scholars like myself, but there may be implications of this work for progressive scholars of color as well. The seamlessness of what Kendall Thomas calls the “jargon” of black racial authenticity, for example, pins down blackness by excluding gayness: “the affirmation of black manhood staged at the Million Man March seemed to *require* the denigration of gay and lesbian African Americans as an enabling condition for its own formation.”⁶ For all of us who, in our writing, seem to speak unproblematically, there is a danger of flattening out, erasing, or suppressing differences that matter, including differences of language, nation, religion, sexuality, gender, and class. Nevertheless, as I will address later, appealing to an unreliable narrator as a way of disrupting our claims to authenticity and authority is not for all academic writing at all times, any more than it has been for all literary writers at all times.

RELIABLE AND UNRELIABLE NARRATORS IN TRADITIONAL PHILOSOPHY

In traditional philosophy, the narrator is not dramatized but conventionalized: like a television news anchor, one learns to speak in the soothingly authoritative tones of a reliable guide. Charles Taylor’s taking up of the demands of “The Politics of Recognition” is a case in point. Because recognition from others is a “vital human need,” Taylor argues, “nonrecognition or misrecognition can inflict harm, can be a form of oppression.”⁷ Invoking a position that he attributes to “some feminists” and “subaltern’ groups,” Taylor casts their demands in a form commensurate with his

own argument. Because they “have been induced to adopt a depreciatory image of themselves,” in Taylor’s bloodless formulation, members of these groups

have internalized a picture of their own inferiority, so that even when some of the objective obstacles to their advancement fall away, they may be incapable of taking advantage of the new opportunities. And beyond this they are condemned to suffer the pain of low self-esteem.⁸

Sidelining structural issues of exploitation and injustice, the essay offers a degree of closure on the vexed issue of cross-difference relations. For Taylor’s purposes, multicultural demands for recognition are evidence of a larger *Zeitgeist*:

Not only contemporary feminism but also race relations and discussions of multiculturalism are undergirded by the premise that the withholding of recognition can be a form of oppression. We may debate whether this factor has been exaggerated, but it is clear that the understanding of identity and authenticity has introduced a new dimension into the politics of equal recognition.⁹

Although the centrality of recognition to modern, liberal democracy organizes the essay from the outset, Taylor’s authoritative voice helps to persuade philosophically disciplined readers that we are being presented with a balanced argument. No doubt creeps into the *text* as to whether Taylor is an adequate spokesperson for indigenous or colonized peoples, blacks, and women.

Taylor performs two important moves that consolidate his position as a reliable narrator. First — paradoxically, given the enormous weight he attributes to being heard and known on one’s own terms — he subsumes all subaltern voices under his own. Although we hear from Jean-Jacques Rousseau often, and in the original French, there is not a single quotation from a gay or lesbian scholar, a feminist, and/or a scholar of color. Second, Taylor enlists disciplinary expectations of philosophical coherence and intelligibility to suppress competing frameworks. The seamlessness of Taylor’s analysis means that readers are not positioned to read between the lines or against the grain, or to wonder if another narrator might tell the story differently. Ignoring arguments and analyses by feminists, lesbians, gays, and people of color that would muddy the clarity of his claims, Taylor’s smooth summary never hints that many members of marginalized and oppressed groups create separate spaces specifically to *escape* from gestures of “recognition” by members of the dominant group.

The fluidity of Taylor’s prose — uninterrupted by different cadences, other emotional registers, alternate projects — and the seamlessness of his organization are artifacts of his position *above* the arguments. His is not an embodied philosophical stance, dwelling within a particular world, nor a relational stance in which the other is expected to give a return. It is dispassionate. Yet Taylor’s own philosophical sensibility infuses the enterprise. Whereas Gloria Anzaldúa or María Lugones might not so much betray or even express passion as specifically call it into play (for it may require courage, deliberation, and discipline to sustain a voice that will be dismissed as subjective or unintelligible), Taylor can wield a highly invested performance of dispassionate argumentation as aesthetic proof of his trustworthiness as a guide through the intricacies of his argument.

By contrast, Jean-Jacques Rousseau's writing persona in *Emile* boldly signals his unreliability. Possibly inadvertently, he invites readers to struggle with the question of how he knows what he claims to know. By constructing a novel in which two incompatible versions of himself serve as characters — the tutor, known as Jean-Jacques, and the narrator, known as Rousseau — he forfeits the possibility of a seamless philosophical narrative. As the tutor, he is suspiciously bland. Patient and self-effacing, Jean-Jacques must appear to be a mere force of nature, for if he had a discernible character Emile might recognize him as having a will of his own; his blandness obscures his manipulation of his young charge. In contrast to the tutor, the narrator, who is known only to the reader, never to Emile, is a forceful character who feels free to scold the reader, announce his paradoxes, and dismiss anyone inclined to quarrel with him. Colorful and highly opinionated, the narrator is modeled on Rousseau but enjoys the added authority of the imaginary “experience” Rousseau claims in his capacity as tutor. The contradictory performances of the tutor and narrator are so rich and complex that readers are given far more material to work with than had we been given a seamless text. Although Rousseau may not intend for us to struggle with the reliability of the narrator's account, he also does not erase its contradictions.

WORKING AGAINST CONVENTIONALIZED TRUST

Among the philosophical devices that bolster a readerly sense of trust are eliminating contradiction, suppressing irrelevance and unintelligibility, using supporting quotations economically, invoking the expected scholarly literature, adopting a recognizable format, telling a familiar story, appealing either to authenticity or to expert authority, and using a calm authorial voice. To work against any of these conventionalized means of promoting readerly trust is to warn readers that the writer is an unreliable narrator.

One way to trouble the seamlessness of academic arguments is to use substantial quotations that exceed our needs as authors. Conventional academic wisdom teaches us to confine our quotations to supporting evidence or counterclaims and to eliminate that which is irrelevant to our project; yet, in eliminating that which is irrelevant to our projects, we may well suppress the ways in which those whom we quote speak against the grain of our inquiry or argument. Other examples of writing against the grain of academic and nonfiction readerly expectations include writing in fragments (perhaps telling a disjointed life story that fails to fit a recognizable pattern); writing in two or more voices that do not merge seamlessly; highlighting ambiguity, contradiction, and uncertainty; interrupting the flow of the argument to raise doubts that are not then laid to rest; using humor to undermine one's authority; and performatively challenging readers' expectations through shifts in genre, voice, perspective, or structure.

A powerful example of a “subjective” writer performatively undermining her own portrayal of coherence and straightforwardness can be found in Ruth Smith's *White Man's Burden*, which includes an account of Smith's work as a white teacher in a Southern school for black girls in the 1930s. Following a chapter that lyrically describes the ways in which “the momentum of our will and our glad surprise at

being there, at knowing each other...inundated every obstacle...or almost did,” Smith turns to another version of the same story, a version that seems to contradict the first.¹⁰ This later chapter is “about loss of heart, about trees without sap, salt without savor, pools without water, Christians honeycombed with fear,” Smith writes.

Of itself it will be a misleading chapter, even as that just finished is... Perhaps they should have been integrated into one picture in the telling, but in that school as in no other place in which I have lived the bright, friendly warmhearted hours were so warm and the terrible, miserable hours were so cold that it seems impossible to talk about them in the same breath.¹¹

Smith’s breach of our expectation of a consistent and uniform picture without contradictions may lead readers back to the earlier chapter where we can see in the words “or almost did” and other hedges the suggestion of important omissions.

A sophisticated example of an “objective” writer performatively undermining her claims to authority can be found in Mica Pollock’s *Colormute*.¹² Depending on the circumstances, Pollock observes, white adults in school either allude to race with casual matter-of-factness or evade race designations altogether. Seeking to trouble her own textual practice of race-naming by rendering it non-matter-of-fact, Pollock variously specifies that a particular race designation is a person’s own customary self-identification, uses hedges like “Chinese-looking” or “Filipino” in quotation marks, narrates her discomfiture at being uncertain as to whether a student is black or Latino, and stresses the students’ own focus on how one is “mixed.” Yet at times Pollock uses “black” matter-of-factly. Only toward the end of the book does she tell us that she came to realize that she rarely flagged “black” as problematic in her field notes. It was as if the “black” designation were self-evident. A few traces of this inconsistency remain in the text, and these seem to me to *perform* her unreliability as a race designator more radically than the diligent use of hedges and explanations. “Filipino-looking” can become a formula after a time, but the contradiction between such a formula and its violation de-authorizes Pollock by unsettling the reader.¹³

As teachers, many progressive scholars seek to de-center our voices, to breach the expectation that our authority should carry the day; our pedagogical attempts at undermining the excessiveness of our intellectual authority may provide some clues for opening up spaces in our writing as well. Actually, I suspect that most of us are not terribly effective at de-centering ourselves pedagogically. It is worth asking ourselves what is at stake beyond what we *say* is at stake when we cling to our various forms of teacherly control, authority, and charisma. What are the costs of our successes as teachers? Just as powerful writers more easily command readerly trust, exemplary teachers may more easily deflect attention away from their powers of control. Yet, with all the limitations in our attempts to de-center ourselves as teacherly authorities, I think we do have some important ways to think about what it means to position ourselves as unreliable educators. Obviously, it is not enough simply to tell our students that they should not defer to our authority. Either they already believe that far more profoundly than we do or they do not believe it at all. To both use and undercut our authoritativeness — which is what it comes down to — we have to provide students with tools, spaces, and resources for thinking along channels other than those that we have cut for them, and we have to provide them

with performative experiences of doing so. Much depends on how we perform our authority: how we listen, how we invite, how we respond, how we take up space, how we make room for students away from our authority, how we read, how we play, how we interrupt the seamlessness of our beautiful syllabi.

We might think similarly about how we wield our authority in educating readers. Do we encourage them to think we have read everything relevant to the topic? Do we seduce them with our nurturing presence? Do we blast them with our brilliance, placate them with our courtesy, reassure them with our logic? Do we make sure we have the last word?

INTERRUPTING PRIVILEGE

Although I am arguing here for a role for unreliable narrators and am speaking in particular to progressive white theorists, for scholars of color the unreliable *reader* may be the more salient part of the writerly transaction. As Robert Stepto observes of Frederick Douglass, “While it was Douglass’s audience’s distrust of him that led to the *Narrative*, it was his increasing distrust of *them* that prompted *My Bondage*.”¹⁴ Because so-called “universal” discourses typically are keyed to the interests and assumptions of those in power, and thus are difficult to adapt to subaltern purposes, members of marginalized or oppressed groups have long had recourse to resistant or hidden discourses. Gossip, conversational codes, slang, trickster tales, folk songs, proverbs, magic, dramatic performance, spirituals, the blues, and humor, for example, all may be vehicles for critiques of the dominant group and have informed a variety of alternative academic formats, including Critical Race Theory counterstories, personal testimonies, and hybrid anthologies of poetry and analysis.

Whereas scholars engaging in trickster and other oppositional discourses take a position outside the dominant academic discourse, using an unreliable narrator persona in academic writing means taking a position inside the discourse but calling into question the assumed relations of trust between readers and writers. Not every philosopher would want or could afford to engage readers in the struggle with her limitations in perspective, experience, assumptions, or writerly craft. How one challenges conventional forms of epistemic and interrogational authority depends in part on the speaker, her project or purposes, and the audience and situation. For progressive white philosophers, the risk in assuming a conventionally objective voice is that we will enhance an authority already invested in race, gender, sex, class, and other power asymmetries, making it difficult for other voices to register in or against our texts without sounding *less* objective, less calm, less rational. A risk in assuming a conventionally subjective academic voice is that we may believe that we know our own story in a way that others cannot, as if prevailing discourses do not help shape the stories we recognize and choose to tell. There is a further risk, too, that in pursuing private and often confessional analyses, we may recenter our own experiences, losing sight of the relational, communal, and structural patterns and projects of which our own story is only a part. Indeed, our stories may do real harm — as when confessions of white racism replay patterns of violence against people of color.

For scholars of color, too, there may be costs in using either the conventionally objective or the conventionally subjective academic voice. Insofar as it represents an approximation of whiteness, the conventional objective voice may impose a borrowed universalism framed in terms of dominant interests and assumptions. Strategically, such a voice may be necessary or valuable, but the cost may be a loss of complexity of engagement. For members of oppressed or marginalized groups, the conventionalized subjective voice allows for important appeals to a kind of experience that outsiders cannot claim already to know. Yet it also risks falling into ready-made purposes and discourses. As Caribbean-Canadian scholar Sherene Razack puts it, “we (people of colour) are always being asked to tell our stories for *your* (white people[’s]) benefit, which you can’t *hear* because of the benefit you derive from hearing them.”¹⁵

There is, of course, no single or perfect solution to our limitations as academic narrators. Any new innovation may introduce some new form of flattening out, seductive clarity, comforting intelligibility. The unreliable narrator is only one device for engaging in more complex writer–reader relationships; it is scarcely an unproblematic option, particularly for those with relatively little professional or political power. There are plenty of tactical reasons *not* to invoke an unreliable narrator in one’s philosophical writing: most obviously, the difficulty of doing so (nicely performed here in a nonperformance), but also the power of an audience not to respond, the pressures of one’s situation (going up for tenure, speaking to an unfamiliar audience), and the need to seize upon whatever tools one may find to get people to listen. Nonetheless, it is worth thinking about our collective assumptions both that we as readers know when to trust a writer’s voice and that we as writers should try to earn our readers’ trust. We might have richer and more generative relationships with readers and writers if they and we know that we can only be trusted so far.

Many thanks to Frank Margonis, Cris Mayo, and Joshua Price for their careful readings and helpful suggestions on the manuscript.

-
1. Jane Austen, *Pride and Prejudice* (1813; repr., Oxford: Oxford University Press, 1932), 3.
 2. Wayne Booth, *The Rhetoric of Fiction*, 2d ed. (Chicago: University of Chicago Press, 1983), 340.
 3. For an analysis of such readerly complicity, see Audrey Thompson, “When Manipulation Is Indispensable to Education: The Moral Work of Fiction,” *Journal of Thought* 32, no. 3 (1997): 27–52.
 4. Liz Stanley, “‘A Referral Was Made’: Behind the Scenes during the Creation of a Social Services Department ‘Elderly’ Statistic,” in *Feminist Praxis: Research, Theory and Epistemology in Feminist Sociology*, ed. Liz Stanley (London: Routledge, 1990), 113.
 5. My point is not that there is nothing outside of the dichotomous objective/subjective construction but that one can rescue only one half of a dichotomy. To rethink objectivity and subjectivity outside of their conventional Western either/or framing, we would have to reinvent them, as indeed a number of scholars have sought to do — as communal, relational, contextual, political, phenomenological, and discursive framings that, while different from one another, are not merely cartoon opposites.
 6. Kendall Thomas, “‘Ain’t Nothin’ Like the Real Thing’: Black Masculinity, Gay Sexuality, and the Jargon of Authenticity,” in *The House that Race Built*, ed. Wahneema Lubiano (New York: Vintage, 1998), 130.

7. Charles Taylor, "The Politics of Recognition," in *Philosophical Arguments* (Cambridge: Harvard University Press, 1995), 226, 225.
8. *Ibid.*, 225. Here Taylor is speaking specifically of women, but he goes on to state that analogous claims are made for blacks and for colonized and indigenous peoples.
9. *Ibid.*, 232.
10. Ruth Smith, *White Man's Burden: A Personal Testament* (New York: Vanguard Press, 1946), 96.
11. *Ibid.*, 112–113.
12. Mica Pollock, *Colormute: Race Talk Dilemmas in an American School* (Princeton: Princeton University Press, 2004).
13. Pollock may or may not have chosen deliberately to leave traces of her changing usage in the text. Either way, the palimpsest effect demonstrates the value of not cleaning up too assiduously after ourselves.
14. Robert B. Stepto, "Afterword: Distrust of the Reader in Afro-American Narratives," in *From Behind the Veil: A Study of Afro-American Narrative*, 2d ed. (Urbana: University of Illinois Press, 1991), 196.
15. Sherene H. Razack, "Storytelling for Social Change," in *Returning the Gaze: Essays on Racism, Feminism and Politics*, ed. Himani Bannerji (Toronto, Ont.: Sister Vision Press, 1993), 92.