FICTION, DEFAMATION, AND FREEDOM OF SPEECH

Collin O'Neil[*]

INTRODUCTION

Speech damages someone's reputation when it leads others to believe that that person has done something that reflects poorly on their character. When that belief is also false the reputational damage is undeserved, and it is the point of American defamation law to protect individuals from suffering such undeserved reputational damage. It is easy to understand why individuals would need protection from false and derogatory claims made about them in works of nonfiction, such as journalism, documentaries, and biographies. But it is not immediately clear why individuals would also need protection from fiction. Although authors of fiction often base their fictional characters on real people, they do not typically make real people characters in their stories. Even when they do put real people in their stories and depict them as doing bad things, the audience is still usually meant only to *imagine* the real people doing those bad things.

Nevertheless, some works of fiction are not only about real people but also do real and undeserved damage to their reputations. It may not be true, as has often been alleged, that Aristophanes's comedy *The Clouds* gave Socrates the reputation for rejecting the gods and corrupting the young that later led to his execution.[1] But readers of parodies of news articles published on sites like *The Onion* and *The Babylon Bee* are sometimes duped, especially when they are already inclined to think poorly of the public figure that is being ridiculed.[2] Of course parodies are believed only when they are not recognized as parodies. But there are other genres of fiction that mix facts into the story, such as biofiction, biopics, and docudramas, and it is not always easy for audiences to distinguish what the author is making up from what the author is, or ought to be, trying to get right.

The biographical drama *Amadeus* suggested that Salieri poisoned Mozart, re-popularizing an old rumor about Salieri that the filmmakers must have at least strongly suspected was false.[3] Salieri, being dead, is in no position to bring a lawsuit. But the villain of the docudrama *When They See Us,* Linda Fairstein, is alive and is

---

[1] *See* Stephen Halliwell, *Did Comedy Kill Socrates?*, OUPBLOG (Dec. 22, 2015), https://blog.oup.com/2015/12/birth-comedy-socrates-aristophanes/.

[2] *See* R. Kelly Garrett et al., *Too Many People Think Satirical News Is Real*, CONVERSATION (Aug. 16, 2019, 8:54 AM), https://theconversation.com/too-many-people-think-satirical-news-is-real-121666.

[3] *See* Alex Ross, *Salieri's Revenge*, NEW YORKER (May 27, 2019), https://www.newyorker.com/magazine/2019/06/03/antonio-salieris-revenge.

suing Netflix and Ava DuVernay, the director, for defamation. Fairstein was chief of Sex Crimes Prosecution during the investigation and prosecution of the "Central Park Five," five Black and Latino teenagers who were convicted of the beating and rape of a jogger in Central Park but who were exonerated years later after a serial rapist whose DNA was found at the scene confessed to the crime and said that he had acted alone. Fairstein alleges that she was defamed in several scenes in the docudrama, including in a scene where she is depicted as concealing potentially exculpatory evidence from the defense and a scene where she is depicted as instructing officers to use harsh interrogation techniques. As a result of her depiction in *When They See Us*, Fairstein's publishing contract was canceled (she had become a best-selling mystery writer since leaving the DA's office), her literary agents dropped her, #cancellindafairstein trended on Twitter, and Glamour magazine expressed regret they had named her Woman of the Year in 1993.[4]

As the docudrama *When They See Us* makes clear, fiction about real people can do serious damage to their reputations. It is another question whether it is ever appropriate to hold an author of fiction legally liable for that damage. One aim of defamation law may be to reflect our pre-legal moral duties of care to avoid damaging others' reputations. If so, one important consideration for determining how defamation law should handle fiction is whether and when an author of fiction would count, morally speaking, as having wrongfully damaged someone's reputation. But defamation law is also answerable to another moral value, namely, freedom of speech, that may be in tension with these pre-legal duties of care. Even when it is plausible that an author of fiction has wrongfully damaged someone's reputation, there might still be a reason of freedom of speech, even an overriding reason, to shield such an author from liability.

This Article will address the question of what limits, if any, freedom of speech would place on holding authors liable for the reputational damage they cause with fiction. By "freedom of speech" I will not be referring to the First Amendment but rather to one conception of the moral idea underlying it. According to this conception, the limits that freedom of speech places on the scope of authors' liability for causing false and defamatory beliefs are whatever limits are necessary to adequately

---

[4] Fairstein v. Netflix, Inc., 553 F. Supp. 3d 48 (S.D.N.Y. 2021). Judge P. Kevin Castel denied in part the defendant's motion to dismiss, finding that several scenes did plausibly carry a defamatory meaning. *Id.* at 58.

protect our interests as potential authors and audiences, and whose costs are acceptable in terms of other interests.[5] To apply this conception, it will be necessary to identify our interests as potential authors of and audiences for fiction about real people, and to assess how these interests would be affected by different limits. Ultimately, I will argue that freedom of speech is consistent with holding authors liable for reputational damage caused by their violations of fiction's "veracity rules" and for reputational damage caused by mistakes that their *target audience* would be expected to make. But liability for beliefs that are traceable to mistakes that only an author's *incidental audience* would be expected to make is, I will argue, prohibited by freedom of speech, so long as the costs of that protection remain acceptable.

I.   DEFAMATION LAW

Although my aim in this article is neither to interpret the current law of defamation in fiction nor the First Amendment, a brief discussion of each will help to prepare the way for the argument to follow. Defamation law applies the same standards to works of fiction as it does to works of nonfiction.[6] One element of defamation is the "of and concerning" requirement, which is satisfied only if the defamatory content is capable of being reasonably understood as referring to the plaintiff. The defamatory content must also be false and provably false—i.e., not an opinion in the legal sense. And the defendant must be shown to have met a certain standard

---

[5] This way of formulating the general framework is closest to T. M. Scanlon's in A Framework for Thinking About Freedom of Speech, and Some of Its Implications (2018) (unpublished manuscript), https://www.law.berkeley.edu/wp-content/uploads/2018/10/Freedom-of-Speech-Berkeley.pdf. Joshua Cohen applies a similar framework in terms of speaker, audience, and third-party interests in *Freedom of Expression*, 22 PHIL. & PUB. AFF. 207 (1993). Seana Valentine Shiffrin develops a framework that emphasizes thinker interests in SPEECH MATTERS: ON LYING, MORALITY, AND THE LAW 79–115 (2014).

[6] Some commentators argue that different standards should apply. Heidi Stam argues that freedom of speech demands categorical immunity from liability for defamation in fiction in Heidi Stam, Comment, *Defamation in Fiction: The Case for Absolute First Amendment Protection*, 29 AM. U. L. REV. 571 (1980). Others stop short of recommending categorical immunity but instead recommend placing extremely demanding conditions on liability for defamation in fiction that very few defendants would satisfy, e.g., that liability should require proof not only that the defendant intended that the audience come to believe the false and defamatory content about the plaintiff but did so from a motive of "classical malice," that is, from ill-will or hatred towards the plaintiff. *See, e.g.*, Dan Rosen & Charles L. Babcock, *Of and Concerning Real People and Writers of Fiction*, 7 HASTINGS COMMC'NS & ENT. L.J. 221, 225 (1985); Matthew Savare, *Falsity, Fault, and Fiction: A New Standard for Defamation in Fiction*, 12 UCLA ENT. L. REV. 129, 161 (2004).

of fault. In the case of public figure plaintiffs like Linda Fairstein, the Supreme Court has determined that the First Amendment requires that they be assigned the burden of proving that the defendant published the false and defamatory statement with "actual malice," that is, with knowledge of its falsehood or in reckless disregard of its falsehood.[7]

Whenever fiction refers to real people by their real names, it will be obvious that the story is "of and concerning" those people. But if liability for defamation of public figures only required, in addition, that the author knew or strongly suspected that the defamatory content was false, this would clearly leave the interests of authors and audiences of fiction inadequately protected. There may be little reason to worry about discouraging authors from publishing nonfiction with defamatory content that they have simply made up or strongly suspect to be false. That's not valuable speech. But we normally both expect and want authors of fiction to make things up when falsehoods would serve the story better than the truth. We want this even when the fiction refers to real people. The recent novel *Rodham* by Curtis Sittenfeld invites the audience to imagine how Hillary Clinton's life would have gone if she and Bill had never married.[8] The fact that nearly every sentence the novel contains about Hillary Clinton is entirely made up does not at all diminish its value as fiction.[9]

But there is another requirement that most courts impose on liability for defamation in fiction, namely, that the publication of the defamatory content "reason-

---

[7] N.Y. Times Co. v. Sullivan, 376 U.S. 254, 279–80 (1964); Gertz v. Robert Welch, Inc., 418 U.S. 323 (1974).

[8] CURTIS SITTENFELD, RODHAM (2020).

[9] For a somewhat skeptical view about the value of falsehoods in fiction, on the grounds that "[h]e or she is breaking the most important rule of our social existence, albeit for purposes that are often important," see Frederick Schauer, *Liars, Novelists, and the Law of Defamation,* 51 BROOK. L. REV. 233, 266 (1985). For the view that protecting the exercise of the imagination is a core First Amendment value, see Jed Rubenfeld, *The Freedom of Imagination: Copyright's Constitutionality*, 112 YALE L.J. 1 (2002).

ably be understood as describing actual facts about the plaintiff or her actual conduct."[10] Courts understand and apply this language in somewhat different ways.[11] But this much is clear. The mere fact that, in a work of fiction, an author depicts a public figure as doing something bad that the author knows they didn't do, because the author made it up, would not be enough on its own to sustain a defamation claim, even if the audience believed it.

I will not say more about how courts handle cases of defamation in fiction since my aims are more theoretical. But it is worth noting that this "reasonably be understood as describing actual facts" condition reflects the law's recognition that, even when an author knows that the defamatory content is false, they should not be held responsible just because their audience believes it. It is my aim in the remainder of this Article to define the proper boundaries of their responsibility, insofar as freedom of speech is concerned.

II. Fiction

Before we can see how our interests would be affected by different limits on liability for defaming someone with fiction, we'll need an account of what fiction is.

The view I will assume here is that fiction-making is the speech act of inviting an audience to *imagine* a proposition.[12] A collection of fictive utterances (in books) or scenes (in movies) whose propositions are meant to be imagined together is a "fiction."[13] Since a *work* of nonfiction can contain one or more fictions (e.g., a phi-

---

[10] Pring v. Penthouse Int'l, Ltd., 695 F.2d 438, 440 (10th Cir. 1982).

[11] For a useful discussion of the way that courts have interpreted the "statement of fact" requirement in connection with fiction, see Mark Arnot, *When Is Fiction Just Fiction? Applying Heightened Threshold Tests to Defamation in Fiction*, 76 FORDHAM L. REV. 1853, 1871–73 (2007).

[12] On a Gricean account, this is elaborated as the act of intending to get an audience to imagine that p in part by revealing to the audience one's intention to get them to imagine that p. Such an account is developed in GREGORY CURRIE, THE NATURE OF FICTION 30–35 (1990), and more recently in KATHLEEN STOCK, ONLY IMAGINE: FICTION, INTERPRETATION AND IMAGINATION 13–44 (2017).

[13] For this definition of a fiction, see STOCK, *supra* note 12, 145–74.

losophy article that contains thought experiments or a memoir that contains myth-ical interludes[14]), the interests of authors and audiences in connection with fiction are not limited to works of fiction.

The idea that fiction-making involves inviting imaginings is widely accepted. But I will make two further assumptions that are more controversial. First, I will take the fictional content of a fiction to be determined by the author's intentions. An author can do a better or worse job of communicating that content and of get-ting the audience to imagine it, but the fictional content itself—i.e., what is *true in the fiction*—is whatever series of propositions the author communicatively intends that the audience imagine.[15]

Secondly, although the content of fictions is paradigmatically invented or made up, so that when that content is true, it is only accidentally true, I will assume that being made up or being non-truth-tracking in some broader sense is *not* a necessary condition for fictionality.[16] An utterance may, on the view I will be adopting, count as fictional even if the author was inviting an audience to imagine that proposition only because the author believed it was true, and even if the author was also, by means of that same utterance, inviting the audience to believe, not merely imagine, that proposition.[17] Whether fiction may or may not be truth-tracking in either of these two senses won't matter much for my purposes, since my primary focus will be on content that the author makes up without intending to deceive. But I will

---

[14] Helena de Bres in ARTFUL TRUTHS: THE PHILOSOPHY OF MEMOIR (2021) cites the example of Maxine Hong Kingston's memoir, THE WOMAN WARRIOR (1976), which switches between memo-ries and myth.

[15] This view is known as "extreme intentionalism" and is defended by STOCK, *supra* note 12, at 13–44. Although I find it plausible, I am adopting it here mainly to simplify the terminology I will use. Readers can substitute their preferred approach.

[16] CURRIE, *supra* note 12, at 42–49, argues that fiction cannot be non-accidentally true. STOCK, *supra* note 12, at 145–74, argues that an invitation to imagine is sufficient for fiction-making. For a range of views on this question, see Fred Kroon & Alberto Voltolini, *Fiction, in* STAN. ENCYCLOPE-DIA PHIL. ARCHIVE (Nov. 12, 2019), https://plato.stanford.edu/archives/win2019/entries/fiction.

[17] Treating an invitation to imagine as sufficient for fictionality would seem to imply that works of narrative nonfiction like many memoirs, some biographies, and so-called "New Journalism," whose content is also meant to be imagined, would count as works of fiction. It won't matter for my purposes whether such works are classified as belonging to a subgenre of fiction that is subject to comprehensive veracity rules or as nonfiction. But for a reason to still count them as nonfiction, see STOCK, *supra* note 12, at 145–74.

need to assume that content in a work of fiction can count as fictional even if it is supposed to be true according to a veracity rule that governs and defines that category of fiction (more on these rules later).

III. REAL PERSON FICTION

Authors of fiction normally select the imaginings they want to produce in their audience with the aim of eliciting further emotional responses such as suspense, surprise, fear, amusement, surprise, and anger. Authors often aim to produce further cognitive effects as well. One common cognitive goal (and perhaps the one most often sought by audiences for fiction) is to convey to the audience *what it is like* to undergo certain experiences or to embrace different values. Fiction can also convey more ordinary facts. Realistic fiction often conveys general facts about the world. Fiction that is "based on a true story" is often a source of biographical information about real people.[18]

Whatever an author's more ultimate goals are, they will have an interest in the freedom to invite their audience to imagine whatever content they think will be most effective for achieving those goals. Inevitably their efforts to achieve those goals will have side-effects on some audience members that they do not intend. Authors' interest in the freedom to create fictional content, then, gives them an interest in the freedom to *disregard* the potential for such unintended effects on at least certain members of their audience. Being potentially liable for causing effects such as false and defamatory beliefs could diminish that freedom.

Why might authors be drawn to write the sort of fiction that could lead their audiences to form such beliefs? It is easy to understand why authors would want to populate their fiction with realistically drawn fictional characters, even though there is a chance that audiences will mistakenly take those fictional characters to be based on a real person if there happen to be enough shared traits. Realism enhances emotional engagement. It is also easy to understand why an author would want to closely model a fictional character after a real person who possesses some intriguing qualities or who has led an extraordinary life.[19]

---

[18] For a full accounting of the potential epistemic benefits of fiction, see Mitchell Green, *Fiction and Epistemic Value: State of the Art*, 62 BRITISH J. AESTHETICS 273 (2022).

[19] For a sense of just how common this practice is, see WILLIAM AMOS, THE ORIGINALS: AN A-Z OF FICTION'S REAL-LIFE CHARACTERS (1985), which pairs thousands of fictional characters with their real-life models.

But it is less obvious what authors and audiences gain from putting real people in stories, especially when the author is planning to make things up about that person. If the real person is interesting and important, why not write a biography? If the problem with biography is that it doesn't give the author the freedom to invent things about the person, then why not base a fictional character on that person instead, which would be even more liberating?

There are some writers and critics who claim that using real people in fiction is artistically (and possibly morally) irresponsible.[20] But there are a variety of compelling reasons to mix real people and fiction in this way, even though it carries a risk of prompting the audience to form false and defamatory beliefs about those people.

IV. Authorial Interests in Real Person Fiction

One reason to put real people in fiction[21] is to set the story against a real background. For the same reason authors often set stories in real cities and have fictional characters participate in real events, they may also want to populate the setting with real people. Leo Tolstoy famously included real generals in *War and Peace.*

For the purpose of contributing to the realism of the story, fictional characters could not substitute for real people.

A harder question is why an author would want to make anything up about a real person if their function is to contribute to the realism of the fiction. One reason

---

[20] For a defense of biofiction against critics such as Ralph Ellison, Virginia Woolf, and Jonathan Dee, see MICHAEL LACKEY, BIOFICTION: AN INTRODUCTION 1–5, 77–97 (2021).

[21] By putting a real person in fiction, I mean using their name to refer to them. If an author uses the name "Ludwig Wittgenstein" with the intention of referring to the same real person as other people who use this name, then the fiction refers to Wittgenstein. If instead the author simply likes this name and intends to use it as the name of a fictional character, then it would not refer to Wittgenstein. For an account of the reference of names in fiction, and a defense of the view that they can semantically refer to real people, see Jeonggyu Lee, *Referential Intentions and Ordinary Names in Fiction*, 180 PHIL. STUD. 1059 (2023). It is also possible to communicate something about a real person by means of fictional content without making that person a character in the story. Satires have real-world targets, and while some satires make their real-world targets characters in the satire, others criticize their targets by means of fictional characters that allude to those people. Romans à clef also allude to real people by means of fictional characters and are intended to imply to audience members "in the know" that the real person actually did some of the things the fictional character is depicted as doing. I won't discuss these more indirect ways of getting audiences to think about real people via fiction.

is to bring the real people in the background into contact with the invented foreground at certain points. In Catherine Lacey's novel *Biography of X,* which purports to be a biography of X written by X's widow, X is given famous friends like Patti Smith and Frank O'Hara, and has an influence on real journalists like Rachel Aviv.[22]

Authors of fiction can also have reasons to place real people at the front and center of the fiction. They may want to tap into audiences' antecedent curiosity about particular real people, or their more general desire to learn about real people who have led interesting lives. A brand-new fictional character cannot satisfy such curiosity.

Still, if the point is to provide information about real people, why not write a biography? Arguably the chief advantage of biofiction over biography is that authors can go beyond the biographical facts and help audiences to imagine that they have direct and complete access to the real person's interior life: their thoughts, their perspectives, and the motivations for their choices.[23] Since there can be no record of their minds, this requires, if not outright invention, at least conjecture. To make this more convincing authors may also want to invent incidents that "explain" why the person came to have the thoughts and motivations they do, as well as incidents beyond those in the record for the purpose of displaying their character.

Even when an audience is already expected to know the biography of a real person backwards and forwards, authors can still take advantage of their interest in that person by inviting them to imagine an alternate history for that person. The novel mentioned earlier—*Rodham*—does exactly that. The audience is invited to imagine how Hillary Clinton's life and political career would have gone (and Bill Clinton's too) if she and Bill had split up in their twenties and never married. Replacing Hillary Clinton with a fictional character would undermine the whole point of this novel, and since it explores a counterfactual situation, the author must make up almost all of the content about her.

---

[22] Dwight Garner, *'Biography of X' Rewrites a Life Story and an American Century*, N.Y. TIMES (Mar. 20, 2023).

[23] Julia Alvarez is an author who explains this reason well. *See* Julia Alvarez, *Fixed Facts and Creative Freedom in the Biographical Novel*, *in* MICHAEL LACKEY, TRUTHFUL FICTIONS: CONVERSATIONS WITH AMERICAN BIOGRAPHICAL NOVELISTS 27–42 (2014). In this interview she also explains why in her biographical novel, IN THE TIME OF THE BUTTERFLIES (1994), she depicts her subject as having slapped the dictator Rafael Trujillo, when in fact she only insulted him verbally.

Authors of fiction whose aim is to ridicule or to criticize a real person will also find it useful to put that real person in the story. Some satires use a fictional character to allude to a real person. But the satirical parodies in *The Onion* and *The Babylon Bee*, which mimic news stories, put the targets of their ridicule in those parodies. Fiction that is critical of real people can also take serious forms, as in biopics like *Raging Bull* and *The Social Network,* and docudramas like *When They See Us*.

V. Weighing Authorial Interests in Real Person Fiction

One reason for putting a real person in fiction and making up defamatory content about them is to deceive the audience and thereby damage that person's reputation. Although this is surely sometimes the motivation behind using real people, such an interest should not be given any weight in determining what is necessary to adequately protect our interests as authors and audiences. We can have legitimate reasons for deceiving others, such as to protect our own or someone else's privacy, or to defend ourselves or others from immediate threats. Authors can even have a legitimate interest in *temporarily* deceiving an audience into thinking badly of someone. In an amicus brief, *The Onion* objected to a requirement that parodies be labeled as parodies, on the grounds that this warning would spoil the enjoyable surprise that occurs when it dawns on an audience that what initially appeared to be a news article is a satire.[24]

But getting an audience of strangers to form a false and defamatory belief about someone—especially when that belief is expected to outlast their engagement with the fiction—would rarely if ever serve any legitimate purpose. It might be argued that even if an author has no legitimate interest in deceiving the audience as such, they might still have a legitimate interest in telling the fictional story that serves as the vehicle for the deception. The interest in telling the story would be legitimate. But holding authors liable for intentional deception leaves them perfectly free to tell that same story without a deceptive intention.

Setting aside intentional deception, then, authors still have a variety of legitimate reasons for putting real people in fiction and depicting them unflatteringly. I think it's safe to say that, if the threat of liability would incentivize authors to re-

---

[24] Brief of *The Onion* as Amicus Curiae in Support of Petitioner at 14–16, Novak v. City of Parma, 2022 WL 5552533 (No. 22-293).

move real people from their fiction and/or refrain from depicting them unflatteringly, then the costs to the interests of authors (and their audiences) could be significant.

But perhaps liability would not have that consequence. Suppose for the sake of argument that taking certain paratextual precautions against misleading audiences, such as labeling the work as a novel, making a disclaimer such as "characters . . . are either products of the author's imagination or are used fictitiously,"[25] or writing a preface that specifies exactly what the author has made up about the real people in the fiction,[26] would be fully effective in preventing the audience from forming any false and defamatory beliefs.

Could authors have a legitimate (and nondeceptive) interest in *not* labeling a work as fiction, in *not* adding a disclaimer, and *not* writing such a preface? Might they even have a legitimate and nondeceptive interest in mislabeling their fiction as nonfiction, or otherwise mimicking nonfiction, or possibly even writing a preface, like the preface to *Robinson Crusoe,* that strenuously insists that the story is nonfiction?[27]

If there were no legitimate interest in forgoing precautions like these, and if they were also entirely effective, then liability for causing false and defamatory beliefs would be unproblematic from a free speech point of view. But neither of these assumptions is correct. Authors can have a legitimate interest in forgoing paratextual precautions, specifically, an interest in enhancing the *immersiveness* of the audience's experience of the fiction.

This is an interest in helping the audience to imagine, not just what the characters are doing, but that they (the audience) are learning about these characters from reading or viewing a work of nonfiction, e.g., a history or a biography or a series of

---

[25] This disclaimer is from MICHAEL BENNETT, YOUNG DONALD (2020), a novel that invites the reader to imagine Donald Trump as a young man inadvertently killing his friend when they were at a military academy and feeling no compunction about it afterwards.

[26] Bruce Duffy's preface in THE WORLD AS I FOUND IT at ix-x (1987) warns the reader that he has Wittgenstein meet Russell one year earlier than he actually did, gave him two sisters instead of three, and made G. E. Moore's marriage occur three years earlier.

[27] Although some readers were duped by Daniel Defoe's THE ADVENTURES OF ROBINSON CRUSOE, James Edwin Mahon, *Novels Never Lie*, 59 BRITISH J. AESTHETICS 323 (2019), argues that Defoe's pretense was not intended to deceive but to indicate that the genre of the novel, which was new at the time, was in closer contact with the facts than the more familiar genre of romance.

letters. For example, *The Biography of X,* mentioned earlier, is not only labeled as a biography, it also includes footnotes citing (fake) sources. These features make it a better prop for the audience's attempt to imagine that they are reading a real biography and thereby facilitate a more immersive experience. Of course, the absence of labels and disclaimers and the presence of mislabeling and mimicry would tend to deceive naive readers and viewers. But so long as the aim is not to deceive, but only to enhance the immersiveness of the experience for savvy audience members, the interest is legitimate.

In any case, labeling and disclaimers are unlikely to be fully effective in preventing the formation of false and defamatory beliefs. As mentioned earlier, a surprising percentage of people are duped by *The Onion's* and *The Babylon Bee's* satirical parodies into believing that politicians have done terrible things. Labeling parodies as parodies (which *The Onion* opposes on artistic grounds) does lower that percentage, but some readers still believe.[28] There is also evidence that even when audience members have been explicitly warned that specific content is invented, and even when they recall that it was invented when asked about it later, they are still more likely to believe it than audience members who were never exposed to that content in the first place.[29]

VI. AUDIENCE INTERESTS

The interests we have as potential audiences for fiction are largely aligned with the interests of authors. The reasons that authors have for putting real people in their stories are also reasons for audiences to want authors to put real people in their stories, and the reasons that authors want to provide a more immersive experience are also reasons for audiences to want them to provide it. Moreover, I argued

---

[28] Garrett et al., *supra* note 2.

[29] Jeffrey J. Strange describes his 1993 study in *How Fictional Tales Wag Real-World Beliefs*, *in* NARRATIVE IMPACT: SOCIAL AND COGNITIVE FOUNDATIONS 263–86 (Melanie C. Green, Jeffrey J. Strange & Timothy C. Brock eds., 2002). Also see Stacie Friend for a good review of the literature in this area in *Believing in Stories*, *in* AESTHETICS AND THE SCIENCES OF MIND 227—48 (Greg Currie et al. eds., 2014). The influence seems to be confined to content that appears to be applicable to the real world, unlike content about fictional characters, and that concerns general facts. But it is possible that fictional content about real people could have a similar influence. If warnings do not always suffice, then to avoid liability, authors may need to modify not just the paratextual framing of their fiction, but the content itself.

earlier that authors have no legitimate interest in intentionally and lastingly deceiving their audiences via fiction. Arguably, audiences have no interest in authors having the freedom to intentionally and lastingly deceive them either.[30]

As we'll see shortly, conflicts of interest can arise between authors and between different segments of their audiences. But let's assume for the moment that all audience members are equally vulnerable to forming false and defamatory beliefs in response to fiction, and that the loss of the fiction that would have been created in the absence of liability would be felt equally by all audience members. If the false beliefs are of a kind that could seriously compromise the audience's welfare or their control over their lives (e.g., false medical beliefs[31]), then it is highly doubtful that it could be in the interests of audiences, on balance, for authors to be free from liability for causing such beliefs. No stories are good enough to compensate for that.

But the beliefs under discussion here—false and defamatory beliefs about public figures—are not the sort of beliefs whose falsehood would put an audience's welfare or autonomy at serious risk. These beliefs can be extremely damaging for the people they are about. These beliefs can also, in the aggregate, have serious consequences for the general public. When political officials and public figures have worse reputations than they deserve, that can make elections unreliable and impair public decision-making. Consequences like these are relevant to determining the proper free speech-based boundaries on liability. But insofar as it is only our interests *as* audiences that are concerned, it is at least conceivable that it could be in our

---

[30] It might seem obvious that audiences have no interest in authors having the freedom to do this either. But the vehicle for the deception—the fiction—may still have value for an audience. Since forbidding an author from using that fiction to deceive leaves the author free to publish that same fiction without that intention, the author has no grounds for complaint. But if an author would be unwilling to create or publish that fiction if they could not use it to deceive, then audiences might be deprived of valuable fiction. See Eugene Volokh, *The Freedom of Speech and Bad Purposes*, 63 UCLA L. Rev. 1366 (2016), for an argument that bad purposes are irrelevant to the value of speech for audiences. However, in the case of fiction, it seems more likely that if an author of fiction were not permitted to use it to deceive, they would not refuse to publish the fiction, but would merely remove the elements of that fiction that were there for no other reason than to accomplish the deception, which might be an overall improvement.

[31] Audiences do acquire many of their beliefs about health care from fictional shows like *ER*. *See* Mollyann Brodie et al., *Communicating Health Information Through the Entertainment Media*, 20 Health Affairs 192 (2001).

interests on balance to accept a higher risk of being led to form such beliefs in exchange for better stories.

In fact, however, the benefits and the costs of liability are not the same for all the members of an author's audience. When we drop this unrealistic assumption and distinguish between what I will call the author's *target audience* and the author's *incidental audience*, we will see that, holding authors liable for causing false and defamatory beliefs would have a very different impact on these different segments of the author's audience.

VII.   TARGET AUDIENCE VERSUS INCIDENTAL AUDIENCE

There will often be a segment of an author's potential audience whose full understanding of the fiction is sought by the author and whose experience of the fiction is prioritized by the author. This is the author's *target audience.* The rest of the audience is the author's *incidental audience.* Authors want their target audience to grasp the fictional content completely. They also want to produce further effects on their audience, such as emotional and possibly cognitive effects, and hope that their target audience will evaluate the fiction positively. Since authors cannot rationally intend to achieve what they don't believe is possible, they must believe that, given what they take the knowledge, abilities, and sensitivities of their target audience to be (i.e., their target audience's level of *competence*), they have provided enough evidence of their intentions in the text or film for the target audience to recover that fictional content and that their audiences are capable of experiencing those emotional and cognitive effects.

In addition, although authors would normally be pleased if the incidental audience understands the fiction as accurately and completely as the target audience does and undergoes the same experiences, the author *prioritizes* the experience of their target audience in the following sense. If there are changes that would guard against misunderstandings by the incidental audience or that would enhance the experience of the incidental audience but would detract from the value of the fiction for the target audience, the author will be disposed against making those changes.

Not all authors have target audiences in this sense. Some authors may be writing for anyone that is linguistically competent and if so, the question of whose understanding or experience should be prioritized would not arise. But most authors do have a target audience. For example, an author may create content that they hope will amuse or entertain a child. They may anticipate that adults (perhaps parents who read it to their children) will not find this content even slightly amusing

or entertaining. But if the author is writing a book for children, they will not be disposed to change it in ways that would make it more interesting for adults if this would detract from its power to amuse children. (In Part XIV I will address the question of how to identify the author's target audience without having to take the author's word for it.)

With this distinction between the target and incidental audiences on board, I will use it to argue that holding authors liable for false and defamatory beliefs that arise from misunderstandings that only the incidental audience would be susceptible to would be contrary to the interests of both the author and their target audience.

VIII.   MISTAKEN INTERPRETATIONS OF FICTIONAL CONTENT

There are different ways in which an incidental audience could misunderstand fictional content about a real person that would lead them to form false and defamatory beliefs about that person. They might think that a character's hyperbolic or ironic criticism of a real person was meant literally. They might also fail to recognize that a character or narrator is unreliable in the story. For example, in *American Hustle*, a film that fictionalizes the Abscam scandal, Rosalyn (the character played by Jennifer Lawrence) claims that Paul Brodeur wrote an article arguing that microwaves sucked the nutrition out of food. Brodeur, a real person, had never made such an argument and sued.

The court struck down Brodeur's lawsuit on the grounds that the audience "would not expect anything Rosalyn says to reflect objective fact."[32] The court may have been right that the audience would have known better than to trust Rosalyn. But if her unreliability had been established with more subtlety, it's possible the audience would have missed it. If the incidental audience also believed (correctly or incorrectly) that the filmmakers would not make anything unflattering up about a real person, they would have been led to form false and defamatory beliefs.

Holding authors liable for mistakes such as these would be contrary to the interests of both authors and their target audiences. To prevent the incidental audience from misinterpreting a character's hyperbolic or ironic statement about a real

---

[32] Jeff D. Gorman, '*American Hustle' Scene Did Not Defame Author*, COURTHOUSE NEWS SERV. (June 30, 2016), https://www.courthousenews.com/american-hustle-scene-did-not-defame-author/.

person as literal, or to make the unreliability of a character more manifest, the author might need to add clarificatory content that would strike the target audience as clumsy or heavy-handed. If that would not be enough to prevent the misunderstanding, they might have to remove the content about the real person entirely. The target audience has nothing to gain from such modifications and something to lose.

That said, the target audience can also be susceptible to mistakes about fictional content. Although the author is, by definition, trying to communicate successfully with the target audience, it is still possible for them to fail in that attempt. The author may have misjudged the level of competence of the target audience. They may lack the background knowledge or abilities that the author was presuming when writing with that audience in mind. Or the author, while correct about their level of competence, may have inadvertently failed to provide as much evidence of their communicative intentions as an audience with that level of competence would need.

But there is an important difference. If an author's *target* audience is prone to mistakes about the fictional content, then it would not be contrary to the interests of the author to make the changes needed to forestall their misinterpretations. Holding authors liable for false and defamatory beliefs based on misunderstandings by the target audience would merely give authors an extra incentive to do what they are already trying to do. (Authors would obviously prefer not to be given that incentive in the form of potential liability. But their interest in avoiding liability as such is not relevant to freedom of speech, only the precautions they would take to avoid it.)

IX. Veracity Rules

Even when the audience is correct about the fictional content, there is another kind of mistake they can make that might also lead to a false and defamatory belief. This is a mistaken application of veracity rules.[33] Subgenres of fiction (e.g., horror, comedy, musical, fantasy, science fiction, biofiction, historical fiction, biopics, docudramas, etc.) are creatures of convention. When an audience recognizes a work as belonging to a certain subgenre of fiction, they form certain expectations about the

---

[33] For more on veracity or accuracy conventions in fiction, see Green, *supra* note 18. On accuracy conventions in art more broadly, see Daisy Dixon, *Lies in Art*, 100 Australasian J. Phil. 25 (2022).

features of that work and are cued to apply certain evaluative standards.[34] Authors rely on these audience expectations in designing their fiction.

Some subgenre conventions support expectations about what the story will be about, e.g., detective fiction will feature a seemingly unsolvable crime. Veracity conventions sustain expectations that certain kinds of content will be true. The veracity rules governing nonfiction subgenres like memoir, biography, and journalism require that *all* the content be true, unless the author clearly signals that some content will be invented and specifies which content that is.[35] These signals can take different forms.[36]

Some subgenres of fiction are partially defined by rules that require *some* of their content to be true. Realistic fiction seems to be governed by a rule that the background setting for the story will be accurately described unless there is a dramatic reason to depart from the truth. Authors of realistic fiction are not usually criticized for making up features of the background when those falsehoods play a role in the story, but they are criticized for inadvertently getting things wrong.

There are also subgenres of fiction with more elaborate veracity rules. Here is one fairly uncontroversial sketch of the rules that govern and define historical fiction, as well as other fiction based on true stories like docudramas, biopics, and biofiction: The main rule is that the content should not contradict the historical record except in minor ways.[37] (It would not count as a violation of this rule, how-

---

[34] On fiction as a genre and subgenres of fiction, see Stacie Friend, *Fiction as a Genre*, 112 PROC. ARISTOTELIAN SOC'Y 179 (2012); Catherine Abell, *Genre, Interpretation and Evaluation*, 115 PROC. ARISTOTELIAN SOC'Y 25 (2015).

[35] Helena de Bres persuasively criticizes several memoirists who claim that these genres give them license to make things up in the service of a better story, so long as the more essential truths are still conveyed. *See* DE BRES, *supra* note 14, at 60–104. Joshua Landy does the same with some documentarians. *See* Joshua Landy, *Don't Feed the Liars!: On Fraudulent Memoirs, and Why They're Bad*, 46 PHIL. & LITERATURE 137 (2022).

[36] Helena de Bres describes various devices that memoirists have used to signal fictionalization. *See* DE BRES, *supra* note 14, at 98. For example, Maxine Hong Kingston's memoir, THE WOMAN WARRIOR (1976), switches to a third-person omniscient point of view when Kingston is making things up.

[37] For a useful discussion of the rules of historical fiction, see Iskra Fileva, *Historical Inaccuracy in Fiction*, 56 AM. PHIL. Q. 155 (2019).

ever, if the author signaled that some content would contradict the record and specified that content.) So long as the record is not contradicted, the rules leave authors largely free to conjecture and to invent in the gaps.

If there is not enough information in the record to turn a real person into a compelling character, they can invent plausible details of that person's biography and endow them with a personality. They can also invent thoughts and conversations, although subject to the constraint that a conversation of that kind might plausibly have occurred at that period. Authors can omit some real people from the story and invent fictional characters. Although fictional characters should not materially change major historical events (the real drivers of events should remain the drivers in the fiction) they can insert them into events as witnesses or participants for narrative purposes.

Authors choose to create works of fiction that belong to one subgenre rather than another, in part, because they intend to make use of the audience's expectations of veracity that come with that subgenre. These expectations can be useful to the author (and audience) in at least three different ways. One way is by enabling the author to *inform* the audience of some facts. These expectations can also be used to guide the audience's interpretation of the fiction. For instance, if the narrator describes an event of which there would probably be a historical record, the audience, familiar with the relevant veracity rules, will expect that description to be true. Thus, if a character later says something inconsistent with that description, the audience will be inclined to regard that character as dishonest or unreliable, or perhaps as speaking ironically, depending on the context. Finally, an audience's expectations of veracity can be useful for getting the audience to shift their evaluative standards. If an unsatisfying ending to a story falls under a veracity rule, the audience will not regard it as a flaw in the fiction as they would if the rules had left the author free to invent a more satisfying ending.

X. Misapplications of Veracity Rules

Even when incidental audiences interpret the fictional content correctly, they can still end up with false and defamatory beliefs because they misapply the veracity conventions. They may mistake fiction for nonfiction and, presuming it is all true, believe fictional content that the author is permitted to invent. This is what is happening to those readers who are duped by the parodies masquerading as news stories published in *The Onion* and *The Babylon Bee*.

The incidental audience might misidentify the subgenre of fiction that the work belongs to and take it to fall under a different subgenre with a different set of veracity rules. One critic worries that the film *Shirley* could be mistaken for a biopic about Shirley Jackson, when it is only loosely based on her life, and that this "may confuse the casual film-watcher."[38] If it was taken as a biopic the audience might come away believing that Shirley Jackson was a "crazy alcoholic" and her husband a "mean critic," as Shirley Jackson's son fears.[39]

Incidental audiences might also recognize the subgenre but be mistaken about the veracity rules that define it. They might think that historical fiction or biofiction is restricted to inventing the interior thoughts and conversations of real people and is not permitted to invent incidents, even when they do not contradict the historical record.

Any of these misapplications of veracity rules could lead the incidental audience to form a false and defamatory belief about a real person if the work contains false and defamatory fictional content.

If authors were liable for false and defamatory beliefs based on misapplications of veracity rules that only the incidental audience would make, this would usually be contrary to the interests of both authors and the target audiences. It could discourage authors of novels like *Biography of X* from trying to facilitate an immersive experience for their target audience and discourage parodists like *The Onion* from trying to surprise their target audience.

Still, there is no guarantee that the presence of paratextual warnings and disclaimers, or the absence of devices designed to give fiction the look and feel of nonfiction, would prevent the incidental audience from misapplying veracity conventions and thereby forming false and defamatory beliefs. If so, then holding authors liable for such beliefs would incentivize them either to stick to the known facts, even when the veracity rules would permit invention, or avoid that subject matter altogether, which would compromise the value of the fiction for the target audience.

---

[38] Elena Sheppard, *New Hulu Film 'Shirley' Isn't a Shirley Jackson Biopic. But It Might Be Better This Way*, NBC News: Think (June 5, 2020, 2:31 PM), https://www.nbcnews.com/think/opinion/new-hulu-film-shirley-isn-t-shirley-jackson-biopic-it-ncna1226326.

[39] Michael Schulman, *Shirley Jackson's Son Talks to His Fictional Mom, Elisabeth Moss*, New Yorker (June 22, 2020), https://www.newyorker.com/magazine/2020/06/29/shirley-jacksons-son-talks-to-his-fictional-mom-elisabeth-moss.

As with mistakes about the fictional content, the target audience, like the incidental audience, can also misapply veracity conventions. The author may have overestimated the competence of their target audience—they may lack familiarity with the subgenre. Or even if the target audience is as competent as the author presumes, the author might still fail to make the intended subgenre sufficiently clear to an audience with that level of competence. But, again, holding an author liable for false and defamatory beliefs that derive from the target audience's mistakes is neither contrary to the interests of the author nor the target audience, since this only incentivizes them to adopt more effective means to their ends.

XI. VIOLATIONS OF VERACITY RULES

I've argued that holding authors liable for false and defamatory beliefs that their target audience forms as a result of either of two kinds of mistakes is not contrary to the interests of those authors. To fully achieve their aims, authors need their target audience to grasp all of the fictional content, to recognize which veracity rules apply, and to form their expectations accordingly. But what if an author wants to violate a veracity rule? Their aim would not necessarily be to deceive their target audience, but merely to write a better story than they could if they stuck to the facts.

Although authors have an interest in getting their target audience to recognize the veracity rules and to expect veracity where appropriate, this doesn't mean that authors must also have an interest in living up to those expectations. Indeed, authors of fiction would seem to have an interest in violating those rules whenever they believe they can improve on reality.

But there is a big problem. Although violating a veracity rule on an occasion could be in the interests of an author and even of an audience, it would not be in the general interests of authors or audiences as such. The ability of authors of fiction to raise expectations of veracity in audiences is very useful to both authors and audiences, for the reasons I described in the previous Part. Authors possess the ability to raise these expectations only because the veracity rules associated with different subgenres of fiction are socially recognized—i.e., because veracity *conventions* exist. Because authors and audiences each as a class benefit from the ability of authors to raise expectations of veracity, authors should comply with these rules, even when

they could create a better story by violating them and perhaps even when there is no risk that the audience would be misled.[40]

Furthermore, legally recognizing the obligation to comply with veracity rules by holding authors liable for violations of veracity conventions when those violations lead audiences to form false and defamatory beliefs would not only not threaten authorial and audience interests; it would help to protect the mutual interests of authors and audiences in the continuing viability of those veracity conventions.[41]

XII.   COMPARING TWO LIABILITY RULES

Here is what has been established so far. First, since veracity conventions are valuable for authors and audiences, liability for false and defamatory beliefs that are caused by violations of veracity rules would not be contrary to the interests of authors or audiences as classes. Second, holding authors liable for false beliefs that derive from misunderstandings or misapplications of veracity rules that only the incidental audience is susceptible to would be contrary to the interests of the author as well as the interests of the author's target audience. Third, holding authors liable for false and defamatory beliefs that were traceable to mistaken interpretations or misapplications of veracity rules by their target audience would not be contrary to the interests of authors or their target audiences.

I have not yet said anything about the interests of authors' incidental audiences. But it seems clear that, if authors were potentially liable for false and defamatory beliefs caused by mistakes the incidental audience would make, this would incentivize them to make changes that would make the fiction more accessible to the incidental audience, which would also reduce the risk of their forming false beliefs.

---

[40] I'm adapting an argument David Owens gives for an obligation of truthfulness in assertion in BOUND BY CONVENTION: OBLIGATION AND SOCIAL RULES 197–215 (2022). The convention of truthfulness in assertion that grounds this obligation is indispensable in a way that the veracity conventions that define subgenres of fiction are not. Testimonial knowledge would not be possible without the convention of truthfulness in assertion; if we lost these veracity-defined subgenres of fiction entirely we would still have nonfiction. But it is significant that people sometimes prefer to learn from fiction when given the choice.

[41] Here I'm adapting an argument in SHIFFRIN, *supra* note 5, at 125–32, in favor of liability for false and insincere assertions of fact (at least in principle, bracketing concerns about error and abuse in the application of the law).

Can we conclude, then, that holding authors liable for false beliefs traceable to mistakes only their incidental audiences would make is good for incidental audiences and bad for authors and their target audiences? In a sense, yes. This would be good for incidental audiences as such, and bad for authors and target audiences as such. But this doesn't yet tell us much about *our* interests as audiences because we can find ourselves in the target audience for some works of fiction and in the incidental audience for others. That is, the same people who make up the target audience for works of fiction aimed at audiences with a certain level of competence will make up the incidental audiences for everything aimed at audiences with a higher level of competence.

It would be better, therefore, in trying to decide what sorts of liability for false and defamatory beliefs would be most consistent with freedom of speech, to consider the interests of audiences at different levels of competence. To this end, I will distinguish two liability rules and examine how audiences at different levels of competence would fare under each of these rules. I will take it as already established that liability for false beliefs caused by intentional deception and by violations of veracity conventions is consistent with freedom of speech, so that the remaining question is whether liability for anything else would be consistent with freedom of speech.

(R1) Liability for only those false and defamatory beliefs that are caused by mistaken interpretations and misapplications of veracity conventions that the author's target audience would make.

(R2) Liability for false beliefs that are caused by mistaken interpretations and misapplications of veracity conventions that anyone, including the incidental audience, would make.

To simplify the discussion to follow, I'll assume that audiences can be classified as having one of three levels of competence (i.e., background knowledge and abilities): the As, Bs, and Cs, where the As have the highest level of competence and the Cs the least.

How will each of these rules affect the interests of authors and the As, Bs, and Cs?

XIII.   TARGET AUDIENCE ONLY (R1) OR INCIDENTAL AUDIENCE TOO (R2)

Let's begin with (R2), which extends liability even to mistakes only the incidental audience would make. If (R2) were the liability rule, this would be contrary

to the interests of authors, since avoiding liability would call for changes that the target audience does not need and would not want. (R2) would mean that authors who were targeting As or Bs would all have to make their fiction fully accessible to Cs.

This "race to the bottom" isn't quite as drastic as it might seem, since it only applies to those aspects of a fiction that have the potential to cause false and defamatory beliefs—that is, the fictional content that is about real people and that attributes (or could be taken to attribute) something defamatory to them. (R2) incentivizes authors to make such content fully accessible to Cs, but the other content in a fiction could still be aimed at a savvier audience.

Nevertheless, (R2) is clearly contrary to the interests of As. Authors that would have otherwise optimized their fiction for As would now make it fully accessible to Cs. As have nothing to gain from (R2) and much to lose.

(R2) would have mixed consequences for Bs. Content that would have only been fully accessible to As would now be rendered accessible to Bs (and Cs). But content that was targeted at Bs would now be rendered fully accessible to Cs, which would diminish its value for Bs.

And (R2) would be good for Cs, who have something to gain and nothing to lose from (R2). Fiction that would have otherwise been fully accessible only to As or to As and Bs would now be fully accessible to Cs. Since there is no audience less competent than Cs, (R2) would not incentivize authors to make any changes out of consideration for a less competent audience that would diminish the value of the fiction for Cs.

Now let's turn to (R1), which limits liability to those mistakes the target audience would be expected to make. (R1) is better than (R2) for authors, since they will not be under any pressure to dumb their fiction down to avoid liability.

(R1) is much better for As than (R2), since authors who were optimizing their fiction for As would continue to do so.

(R1) has mixed consequences for Bs. Fiction targeting As will not be made fully accessible to Bs. But fiction targeting Bs will not have to be watered down for Cs.

For Cs, (R1) is much worse than (R2). (R1), unlike (R2), does not encourage authors writing for As or Bs to make their content fully accessible to Cs. And since there is no audience less competent than Cs, they have nothing to gain from (R1).

Where does this leave us? Although we can say that (R1) better serves our interests as authors, we cannot say, without qualification, that either (R1) or (R2) better serves our interests as audiences. (R1) is better for the As. (R2) is better for the Cs. And for the Bs, (R1) is better than (R2) in one respect and worse than (R2) in another.

Nevertheless, I think there is a strong case to be made in favor of (R1). (R2) is better for Cs in part because it encourages authors who would have otherwise optimized their fiction for As or Bs to make it fully accessible to Cs. But there is no guarantee that these authors would prefer writing for Cs rather than write nothing at all. Furthermore, even under (R1), there would still be plenty of fiction fully accessible to Cs, since a majority of fiction is already aimed at the broadest possible audience. Under (R2), by contrast, the relatively little fiction that would have been optimized for As and for As and Bs is reduced from little to nothing.

Secondly, while (R2) helps to prevent Cs from forming false and defamatory beliefs by making all fiction fully accessible to them, under (R1), Cs would still have forms of self-help available to them to manage fiction aimed at As and Bs. For one thing, the Cs could improve their level of competence and would be encouraged to do so under (R1), whereas under (R2), they would have no incentive to do so.[42]

Thirdly, while incidental audiences are at higher risk of forming false beliefs under (R2), anyone who is especially concerned to avoid forming such false beliefs still has the option of looking things up online. There is evidence that readers who go online to check the facts are much less susceptible to acquiring false beliefs from fiction.[43]

So, while there are things that Cs can do to make their situation under (R1) just as good or better than it would be under (R2), there is nothing As and Bs can do under (R2) to make their situation as good as it would have been under (R1). Therefore, I think (R1) does a better job overall of protecting authorial and audience interests than (R2).

---

[42] For a related argument that liability for the expected responses of audiences rather than only the reasonable responses would result in a "dumbing down" of the discourse, and also that restricting liability to the reasonable responses would give the less competent audiences an incentive to improve, see Lyrissa Barnett Lidsky, *Nobody's Fools: The Rational Audience as First Amendment Ideal*, 2010 U. ILL. L. REV. 799 (2010).

[43] Amalia M. Donovan & David N. Rapp, *Look It Up: Online Search Reduces the Problematic Effects of Exposures to Inaccuracies*, 48 MEMORY & COGNITION 1128 (2020).

Even if this is correct, I am not yet in a position to conclude that (R1) is required by freedom of speech. The reason is that, under the model I'm deploying, this tighter limit is required by freedom of speech only if the increased reputational damage that would be expected under (R1) compared to (R2) is still *acceptable*. I won't take a position on the acceptability of these additional costs, and so my conclusion will have to be conditional: *If* those reputational costs remain acceptable, then freedom of speech requires (R1).

XIV.   APPLICATION TO *WHEN THEY SEE US*

I've argued that authorial and audience interests are best served by (R1), which immunizes authors from liability for causing false and defamatory beliefs that result from mistakes that only their incidental audience would be expected to make. But freedom of speech is consistent with holding authors liable for false and defamatory beliefs caused in the following ways: (a) intentional deception, (b) violations of veracity conventions, (c) mistakes of interpretation that even their target audience would be expected to make, and (d) misapplications of veracity rules that their target audience would be expected to make.

How would these conclusions bear on the question of whether the filmmakers of *When They See Us* may appropriately be held liable for defaming Linda Fairstein? I don't know whether the filmmakers made up the scene that depicted Linda Fairstein as concealing evidence, but let's suppose for the sake of argument that they did.

The first question is whether the filmmakers intentionally deceived the audience. It is clear from their public comments that the filmmakers did intend the audience to believe at least some of the content that depicted Fairstein negatively, since they wanted to hold her and others accountable.[44] But that doesn't guarantee that they intended that the audience believe the scene that depicted Fairstein as concealing evidence.

The next question is whether the filmmakers violated a veracity rule of docudrama. In the case of docudrama, there is arguably a veracity rule to the effect that the content will not contradict the public record except when such a contradiction is clearly and specifically signaled. The filmmakers did run a disclaimer with the

---

[44] The director Ava DuVernay was quoted as saying "And so the goal of this—okay, Linda Fairstein, okay Elizabeth Lederer, okay all of these people on this particular case we need to be held accountable." *Fairstein v. Netflix, Inc.*, 553 F. Supp. 3d 48, 60 (S.D.N.Y. 2021).

credits that says that while the film was "inspired" by actual events, "certain characters, incidents, locations, dialogue and names are fictionalized for the purposes of dramatization."[45] But this is a very general disclaimer. Only a disclaimer that specified the invented content would avoid violating the veracity rule. There are alternative ways to specify which content is invented, e.g., in the biopic *I, Tonya*, the characters occasionally break the fourth wall after a false scene to say "This did not happen."[46] But it must be specified in some way or other.[47]

Even if Fairstein did not conceal evidence, let's assume that depicting her as having done so does not, strictly speaking, contradict the public record. The filmmakers might then claim that they were merely inventing in the gaps. But it's possible there are some other veracity rules governing docudrama that restrict what can be invented in the gaps. One proposal is this: "[A]ny dramatic overlays should not create a more defamatory impact than the sources themselves."[48]

Such a rule could be understood in two ways. On a strict interpretation, docudrama should not invent any scenes that depict Fairstein as behaving badly. On a looser interpretation, it should not invent any scenes that depict her as doing anything *worse* than what she is known to have done. The looser rule would not permit audiences to trust that any unfavorable depiction must be true. But it would permit them to trust that, even if they believed everything the docudrama depicted Fairstein as doing, they would not end up with a lower opinion of her character than she deserved.

The filmmakers would have violated this rule on either interpretation, since depicting her as concealing evidence is depicting her as doing something bad, as well as something worse than what she is known to have done. Perhaps this is not a genuine veracity rule for docudrama. Even so, it's worth considering whether it

---

[45] *Id.* at 60.

[46] Jordan Crucchiola, *A Fact-Checked Guide to* I, Tonya, VULTURE (Dec. 11, 2017), https://www.vulture.com/2017/12/a-fact-checked-guide-whats-true-and-whats-not-in-i-tonya.html.

[47] The reason for this is that the veracity conventions would no longer sustain expectations if general disclaimers could cancel veracity rules. The audience has no way of knowing which content not to trust, and will tend to mistrust it all. Specific disclaimers do not have this problem.

[48] Amy J. Field, *A Curtain Call for Docudrama-Defamation Actions: A Clear Standard Takes a Bow*, 8 LOY. L.A. ENT. L. REV. 113, 125 (1988). Field distills this rule from the judgment in Davis v. Costa-Gavras, 619 F. Supp. 1372 (S.D.N.Y. 1985).

ought to be, since defamation law can not only enforce existing veracity rules but help to establish them as well.

Finally, even if *When They See Us* did not violate any veracity rules, the filmmakers would still appropriately be held liable if there was a mistake about the fictional content or about the proper application of veracity rules that the target audience would be expected to make that would have led them to form the belief that Fairstein concealed evidence. Since it was true in the fiction that Fairstein concealed evidence, the audience would not be making a mistake about the fictional content. But possibly they would have misapplied the veracity rules.

The target audience would not have mistaken *When They See Us* for a documentary, since it has famous actors and a disclaimer. But the filmmakers did publicly describe the docudrama as "100% real" and used slogans such as "a piece of American History gets rewritten with the truth" and "you can't argue with facts,"[49] which might have led the target audience to expect the show to display a much more thoroughgoing fidelity to the truth than the usual docudrama.

The filmmakers might reply that their target audience was actually an audience that was already thoroughly familiar with all of the facts of the case and so would have recognized that their claims that the content was "all true" were not to be taken seriously. This is implausible, but it raises the important question of how the law could identify an author's target audience without having to rely entirely on the author's word.

The answer relies on the definition of an author's target audience. If an author has a target audience, then there should be indications that the experience of this audience has been prioritized over the understanding and experience of the rest of the author's potential audience (i.e., the incidental audience).

Suppose we want to determine whether an author's characterization of their target audience is sincere. One form of evidence that would support the author's claim would be the presence of elements in the fiction that the putative target audience would be expected to catch (e.g., cultural references) but that would be expected to go over the heads of the rest of the audience.

Evidence that would call the author's claim into question would be the presence of elements that the putative target audience would not be expected to need or want,

---

[49] Fairstein v. Netflix, Inc., 553 F. Supp. 3d 48 (S.D.N.Y. 2021).

but that would improve the putative incidental audience's understanding or enjoyment. In the case of *When They See Us*, it could not be plausibly claimed that the filmmaker's target audience was an audience that was already thoroughly familiar with the facts of the case, since such an audience would not have needed or wanted as much exposition as the film provided.

Compare this with another recent defamation case.[50] In the Netflix show *Queen's Gambit*, the announcer of the final chess match says, regarding the fictional character Beth Harmon, "the only unusual thing about her really is her sex, and even that is not unique in Russia. There's Nona Gaprindashvili, but she's the female world champion and has never faced men." Gaprindashvili, who is a real chess master, had actually competed against 59 men. (A court denied Netflix's motion to dismiss, and Netflix ultimately settled.[51]) In this case, I think it is plausible that the filmmaker's target audience would not have formed the false belief about Gaprindashvili. There is no reason to refer to the real Gaprindashvili rather than a fictional character except to appeal to chess aficionados, who presumably would have known that she competed against men. So it is possible that it was only the incidental audience that would have been vulnerable to forming this false and defamatory belief, in which case liability for causing that belief would violate the free speech-based restrictions on liability that I have defended in this Article.

---

[50] Gaprindashvili v. Netflix, Inc., No. 2:21-cv-07408-VAP-SKx, 2022 WL 363537 (C.D. Cal. Jan. 27, 2022).

[51] Matt Stevens, *A Chess Champion Resolves Her 'Queen's Gambit' Lawsuit Against Netflix*, N.Y. TIMES (Sept. 6, 2022), https://www.nytimes.com/2022/09/06/arts/television/netflix-queens-gambit-nona-gaprindashvili.html.