



PROJECT MUSE®

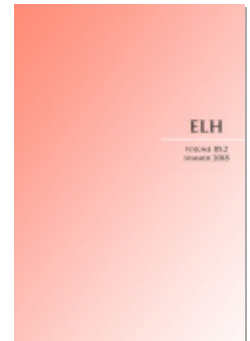
Why Literary Time is Measured in Minutes

Ted Underwood

ELH, Volume 85, Number 2, Summer 2018, pp. 341-365 (Article)

Published by Johns Hopkins University Press

DOI: <https://doi.org/10.1353/elh.2018.0013>



➔ *For additional information about this article*

<https://muse.jhu.edu/article/696232>

WHY LITERARY TIME IS MEASURED IN MINUTES

BY TED UNDERWOOD

Seventeen years ago, in an essay titled “Formalism and Time,” Catherine Gallagher argued that critics are bad at understanding narrative form as something that *takes* time.¹ Instead we try to convert narrative into a timeless structure, or condense stories into a few scenes that convey the meaning of the whole. Whether it’s Jane Eyre walking back and forth on the third story of Thornfield, or Gabriel Conroy watching the snow fall outside his window, we understand fiction by identifying moments of heightened significance. These could be epiphanies or anticlimaxes. In Gallagher’s view the value of these scenes depends less on their specific content than on their rhetorical function, which is to reconcile time with timelessness. She sees critical tradition as deeply shaped by Walter Pater’s dream of cheating death by embracing ephemerality in the form of a single “hard gem-like” moment that, paradoxically, becomes eternal.² A moving aspiration, but also, according to Gallagher, a way of undervaluing the dailiness of life, and long Victorian novels.

This would be an interesting argument under any circumstances, but it’s a particularly remarkable thing for Gallagher to have written in the year 2000, when she was also collaborating with Stephen Greenblatt on a theoretical defense of New Historicism. After all, the New Historicist critic does for historical time exactly what Gallagher’s Paterian critic does for narrative—that is, condense it into a brief scene (an anecdote) that crystallizes the meaning latent in a larger mass of events. This mode of condensation didn’t flourish in New Historicism by happenstance. Gallagher and Greenblatt explicitly theorize the “effect of compression” as an appropriately literary approach to history.³ The anecdote becomes for historical narrative what the detail is for literary realism, conveying Erich Auerbach’s “confidence that in any random fragment plucked from the course of a life at any time the totality of its fate is contained and can be portrayed.”⁴ While Gallagher’s essay diagnoses temporal condensation as an attempt to evade mortality, her coauthored book presents it as a necessary principle of historical understanding, producing “a touch of the real” that disrupts the “generalizable typicality” of the “Big Stories” told by Marxist or Annaliste historians.⁵

I don't mean to criticize Gallagher for exploring both sides of this issue. Temporal condensation is a rhetorical move, not a policy proposal: it may be less important to reach a stable judgment about it than to understand its centrality to literary criticism. This centrality has not been expressed only through New Historical anecdotes and Auerbachian fragments. Literary scholars' titles are often similarly organized by an implicit tableau: *The Madwoman in the Attic*, "The Halted Traveler," *Learning to Curse*.⁶ All of these phrases evoke a brief episode from which we can unfold a larger structure of feeling. That leap across scales of time—connecting collective history to a moment of individual experience, and lending immediacy to the past—is one of the distinctive strengths of literary criticism.⁷

However, the assumptions underlying this gesture are far from self-evident. Why are short spans of time so central to our discipline? Novels commonly cover twenty or thirty years. In some subgenres (science fiction, James Michener's epics) it is not unusual to range across centuries. Why is experience measured in seconds or minutes more appropriately literary than experience measured in weeks or months?

The question becomes urgent for me because much of my own literary research explores long timelines. This is a self-interested choice, not a normative stance. I don't believe that large scales of analysis are more important than resonant details. I just find that, in practice, century-spanning questions tend to be worth investigating, because literary scholars have often left questions on that scale unexplored, or at least unresolved. But these questions have been left unresolved, of course, because they don't fit our discipline's rhetorical templates. Where large historical questions could be condensed into a single case study, someone else has often already done it. The live opportunities I discover tend to be located in aspects of history that don't condense easily—gradual, sprawling trends that can't be represented synecdochically, because their significance depends on the relative position of many different examples.

Gradual and sprawling kinds of change needn't pose a problem for literary research if our discipline is willing to adopt different evidentiary and rhetorical strategies for different scales of analysis. But that is often a polarizing suggestion. For many literary scholars, a particular, fine-grained scale of evidence is bound up with the cultural mission of literary study. We have already seen *Practicing New Historicism*, for instance, champion anecdotes against historians' *grands récits*. This article will end by challenging the assumption that any scale of analysis is uniquely appropriate for literature, but I would like to start

by understanding the assumption. Why do we feel that brief experiences are distinctively literary?

In “Formalism and Time,” Gallagher diagnoses critics’ reliance on small units of time as an evasion of mortality, but she doesn’t tell us where the idea came from. The one clue we might glean is that her story begins with fleeting moments of inspiration in Percy Bysshe Shelley’s “Defence of Poetry.” This Romantic text may remind us of James Chandler’s *England in 1819*, or William Wordsworth’s halted traveler, used by Geoffrey Hartman to define the Romantic lyric as “a meditative slowing of time.”⁸ But it should also remind us of something broader than Romanticism: the rise of the lyric as the paradigmatic poetic form. In 1848, when Edgar Allan Poe writes that “a long poem” is a “contradiction in terms,” we have already taken a big step toward identifying literature with brief experiences.⁹

But the rise of lyric poetry can’t be the whole story. By the middle of the nineteenth century, critics are already taking the novel seriously, so it doesn’t seem inevitable that the lyric should have defined the granularity of time for all literature. No, the decisive move was the choice to understand narrative, too, through representative episodes and moments of lyrical insight. It is a choice made explicit in the structure of *Mimesis*, where Auerbach converts three thousand years of the history of narrative into a sequence of fragmentary scenes.

So how did we reach a point where it seemed natural to understand a whole novel by focusing on five minutes of the story? How did that become not only a valid approach, but—for many scholars—the appropriately literary one?

I: THE COMPRESSION OF FICTIONAL TIME

One plausible lead comes from Auerbach, who quietly acknowledges that his own reliance on fragments echoes the fragmentation of time in recent modernist fiction.¹⁰ Literary historians are broadly in agreement that early twentieth-century fiction had in fact developed new ways to represent time by juxtaposing lyrically compressed episodes.¹¹ Gérard Genette credits the innovation specifically to Marcel Proust. Before Proust, he says, novelists alternated between passages of “scene” and “summary.”¹² In a scene, the time we spend reading may be loosely equivalent to the time represented in the world of the story; events happen, as it were, close up and in slow motion. A passage of summary, on the other hand, will quickly traverse weeks or months where our protagonist has an illness or learns to love the students in her rustic

schoolhouse. According to both Auerbach and Genette, Proust's innovation is to cut out the summary and simply juxtapose one scene with another.¹³ Genette characterizes the time that passes between these scenes as "ellipsis": time that must have passed in the story but isn't narrated in the text.¹⁴ Replacing summary with ellipsis makes fiction, in effect, all scene: all slow motion. Narratologists who aren't French haven't always given Proust personally this much credit, but they tend to agree that the shift toward scene was a modernist innovation.¹⁵

And without distant reading, here is where the article would end, because we have a plausible story. The compression of time that began to define poetry in the Romantic era spread to fiction later, when modernists slowed narration to a phenomenological pace. Gallagher and Auerbach both hint at this account. It is a good fit for several examples we are likely to remember from Proust, Virginia Woolf, and James Joyce. It could also be a polemically useful story—if we wanted to blame our obsession with moments, for instance, on the dubious aesthetics of high modernism. In fact, this is a great story in every way, as long as we don't test our memory of three or four examples by gathering a hundred books and comparing them in detail.

How would we do that? Prompted by Genette, my first thought was, perhaps we should just measure how much space on the page different writers allocate to scene or summary. But, of course, dividing the two is going to be hard. How much time has to pass in a scene before it becomes summary? The dividing line could shift in different books. And, actually, we have no reason to assume that this is a binary division at all. The binary contrast between scene and summary is a critical convention unsupported by much solid evidence. So here is where numbers start to become useful. We don't have to divide passages into binary categories; instead, we can start by treating this as a continuum, and characterize the amount of time that elapses in a given passage—whether that's ten minutes or a month. Later we can look at all the passages and ask whether the contrast between passages of different lengths gives us evidence for a binary division.

Our description of elapsed time doesn't need to be exact. Fictional time isn't exact: how long is a remembered afternoon? We also don't need to be objective. Different people estimate time differently. I collaborated with a couple of graduate student colleagues (Sabrina Lee and Jessica Mercado) explicitly to acknowledge subjectivity. We divided ninety novels between us, and covered almost three hundred years. Some novels were chosen entirely randomly from a large digital library. Others were chosen because they were bestsellers or prominent

in the academic canon: we wanted to make it possible to compare different subsets of the literary field. From each novel we characterized 16 passages of roughly 250 words each (“roughly” because we divided at paragraph and sentence boundaries when possible.) Four of the 16 passages were always the first 500 and last 500 words of each novel, because I was curious about the temporal zooming in or out that might happen there. We allowed a computer to select the other twelve passages randomly from the middle of the narrative.

Then we tried to say how much time is narrated in each passage. We aimed to capture the duration described in the plot—diegetic time, or what the Russian formalists called *syuzhet*—rather than a linear chronology of underlying events. We also anticipated a range of familiar paradoxes. For instance, a break in narration could be recorded as time that passes in ellipsis. At first, we also had a separate category for subjective time, to acknowledge memory and anticipation. But as we got a few novels into the process it became clear that the boundary between subjective and objective time is really a question about scale. For instance, suppose you meet the monster in *Frankenstein* (1818) and he begins to tell his tale. Is narrated duration now the hour or so it takes him to talk, or the month he’s describing? We decided it’s the month. If the edges of his story fall outside our frame of 250 words, we’re inside his narrative. On the other hand, if a character says something about her whole childhood in a short passage of dialogue, and we can see the edges of that subjective perspective, we count it as the minute or two the dialogue would take to speak. In short, we recorded what time looks like at a particular textual scale—250 words, a little less than a page. Since duration depends on scale, it makes no sense to estimate the duration of the whole narrative by multiplying: as if to say, *an hour a page multiplied by two hundred pages equals X*. Narrative is by no means that linear. A different scale of measurement would give us a slightly different picture; nothing prevents scholars from measuring pace in many different ways.

So what should we expect to see at the scale of a page? The account we get in Genette is that the alternation of scene and summary remained fairly stable “up to the end of the nineteenth century.”¹⁶ So if we take 16 passages scattered across each novel, and average them to find out how much fictional time typically passes in 250 words, we might expect to see a steady pace somewhere between scene and summary: shall we say, roughly six hours a page?

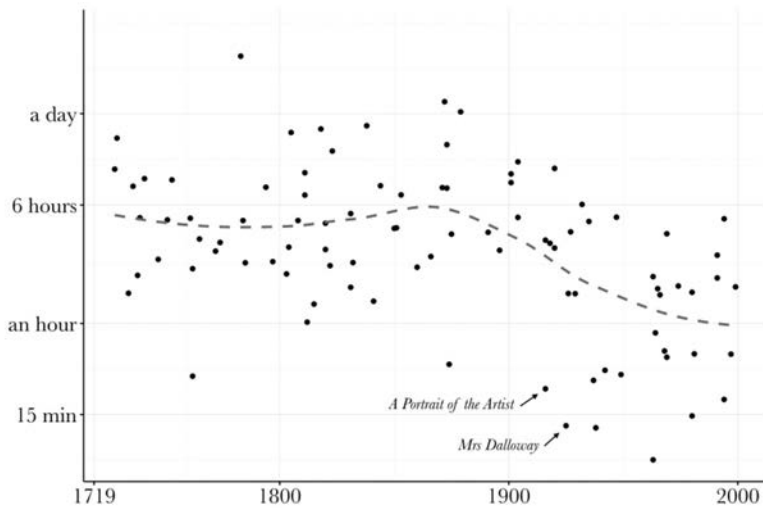


Figure 1. In a hypothetical world inferred from narratological criticism (not real data), this image describes the average length of time described in 250 words of narration. The vertical axis is a logarithmic scale; each dot represents a single work of fiction.

Figure 1 represents a hypothetical world of that kind. On the vertical axis is the amount of time narrated, on average, in 250 words. It's a logarithmic scale, so the distance between 15 minutes and an hour is the same as the distance between 6 hours and a day; in both cases we multiply by four. The average pace of narration remains more or less stable up to the end of the nineteenth century. Then Proust cuts out the summary. Now fiction is all scene. Diegetic time gets slower, more phenomenological: there's going to be less time on each page. At first it's just a few experimental writers doing this in English—Woolf and Joyce. But then eventually their example diffuses outward, and the pace also drops toward scene in the rest of literature. Perhaps in the middle of the twentieth century this effect is further amplified by the influence of film and television, which (like most forms of drama) tend to equate represented time with the time of representation.

I'm not arguing that this picture is exactly what you, personally, would expect to see. Critical tradition hasn't posed this question yet with the degree of precision implied by Figure 1. So we don't really have clear, shared expectations. But I guarantee that once the actual pattern is revealed, it will seem as though we might have guessed

whatever it is we do see. A phenomenon called hindsight bias makes it easy to conclude retrospectively that we already knew, or could easily have inferred, the answer to a problem. In distant reading this effect is particularly powerful, because we are forcing careful answers to large-scale questions where the discipline has taught itself to be content with charismatic guesses. (Genette's off-hand confidence that Proust was responsible for changes in narrative pace is a typical instance.) So before revealing any actual evidence, I thought I would share one thing we might have guessed if we relied on existing accounts of the history of narrative pace drawn from influential scholars like Auerbach, Genette, and Seymour Chatman.

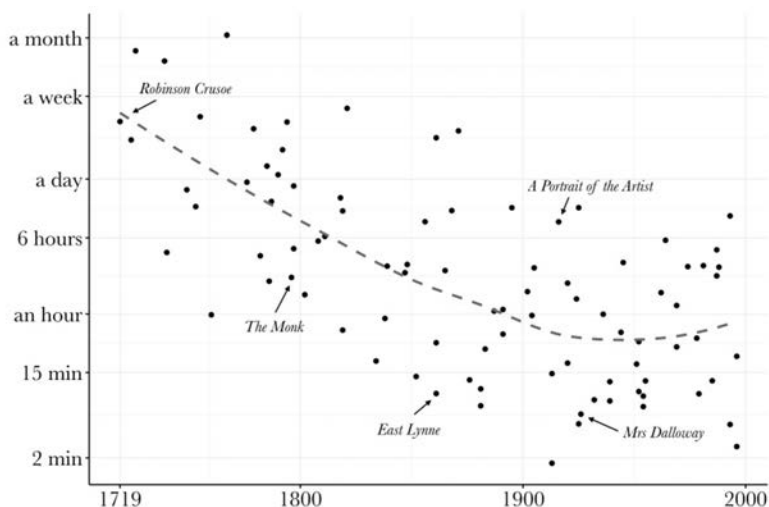


Figure 2. The average length of time described in 250 words of narration (actual data). The vertical axis is a logarithmic scale. The average is calculated as $\text{mean}(\log(\text{passage_duration}))$; the sixteen passages sampled from each volume are weighted equally.

Figure 2 is the pattern we actually see. The pace of narration is already dropping rapidly in the eighteenth century, and it continues to drop to the early twentieth. Then there is perhaps a subtle movement in the other direction, or at least a leveling-off. This picture raises doubts about our first guess that Romantic lyric initiated the compression of time. And it is almost the opposite of the story initiated by modernism

we might have extracted from Genette. Far from being the moment when a new slower pace was introduced, the early twentieth century is the period when the slowing stops, or even reverses.

But debate about modernism is not what I want to dwell on. More interesting: the change between the eighteenth century and the twentieth is enormous. The vertical axis is deceptively compressed here, because it's a logarithmic scale. But fiction goes from covering several days on each page, to covering roughly thirty minutes. That's a hundredfold compression of time. It must be one of the biggest, simplest changes in the history of fiction, bound up (as cause or effect) with many other things. At the scale of thirty minutes you're simply going to narrate different aspects of human life than you would narrate at the scale of two days.

I predicted it would seem that we must have known this already. A trend like this seems too big to hide. All the evidence is in the open; the volumes used in our study are mostly well-known books. Many of us have read, say, half of them. How could we not already know this?

Figure 2 makes one key reason obvious: while there is an enormous difference between the eighteenth-century mean and the twentieth-century mean, the variation around the mean in each century is also enormous. A pattern like this would be very difficult to infer from memory of eight or ten books; you could easily happen to select a group of examples that would show little change across the timeline. And of course, we are also averaging 16 passages from each title in order to infer the average pace for each book. If we illustrated pace in the ordinary way, by quoting a passage or two, it would be very difficult to reach any conclusion. The pace of narration varies too much inside each volume.

In fact, it would be possible to respond to Figure 2 by asking whether changes in the mean pace of narration even matter. If variation around a mean is so large that the mean cannot be grasped intuitively by readers, does the mean even have literary significance? That is a valid question, and I'll give it detailed examination in a moment. But first I want to point out that distant readers are presented with two alternative critiques that cannot be advanced at the same time. On the one hand, we commonly confront the objection that our results are too transparent: they are things a reader might have guessed intuitively from diffuse recollection. If that turns out not to be true, we immediately confront the opposite objection: any pattern that isn't transparently legible in a reader's memory is rejected, as too subtle to matter.

If any evidence that evades the first of these objections is necessarily subject to the second, then distant readers are confronting a form of resistance that cannot be changed by evidence at all. Rather, this pair of objections seems to express a definitional assumption: literary scholarship is simply defined as the reinterpretation of patterns that were already legible in a reader's memory. Any evidence that might have escaped our attention at the scale of individual reading is dismissed in advance. If we accept this argument, distant reading (and many other kinds of research) become impossible.

For literary research to be possible, there must be some ambiguous space between patterns that are transparently legible in our memories and patterns that are too diffuse to matter. In fact this ambiguous space is large and important. We often dimly intuit literary-historical patterns without being able to describe them well or place them precisely on a timeline. For instance, students may say that they like contemporary fiction because it has more action than older books. I suspect that changes in pacing are part of what they mean. There is actually plenty of violence in *Robinson Crusoe* (1719), but it tends to be described from a distance, in summaries that cover an hour or two. We don't see Crusoe's fingers slipping, one by one, off the edge of a cliff. Twentieth-century fiction is closer to the pace of dramatic presentation. Protagonists hold their breath; their heartbeat accelerates; time seems to stand still. This pace may feel more like action, or even (paradoxically) faster, although diegetic time is actually passing more slowly from one page to the next. Even high-school students can feel this difference, although they may not describe it well. Narratologists have described it well, but typically credit it, mistakenly, to modernism. This change is a real literary phenomenon—in fact a huge one. But to trace its history accurately we need numbers.

Problems like this one are the motivation for distant reading. I have borrowed Franco Moretti's apt phrase for this project, but I want to quickly brush away several misunderstandings that have accreted around it. Distant reading doesn't have to involve so-called big data. We are working here with ninety books. Nor does the value of distant reading depend on the premise that we care equally about canonical and obscure volumes. You will recall that we selected these ninety books in several different ways. If we contrast the canonical works and bestsellers to the others that were selected randomly, it turns out that there is no consistent difference between the three sets: the trend at stake here seems to affect literary practice as a whole. We fail to grasp the shape of the trend intuitively, not because our reading has been

limited to the canon, but simply because our memories aren't as good as we like to pretend.¹⁷ We may have read these books, but we can't remember them—all at once—well enough to compare them. We may dimly realize that a change happened, but to trace the change we will have to supplement our memories, and use numbers to measure differences of degree.

II: THE ADVANTAGES AND RISKS OF NUMBERS

Supplementing the human memory does not have to be a high-tech project. The works I take as models for distant reading—say, Janice Radway's *Reading the Romance*—generally rely on numbers, but not always on computers.¹⁸ In this project, too, we could have recorded our responses on index cards, and used a slide rule for the logarithms. I prefer to call this mode of inquiry distant reading—rather than digital humanities—because it is fundamentally about a new scale of historical inquiry, not about digital technology.¹⁹

Of course, computers do make new things possible. This article can be relatively low-tech because it stands at one remove from the text, characterizing literature through the responses of readers. But if we wanted to grapple more closely with details of literary language across thousands of books, new methods drawn from machine learning might become useful. Machine learning is a real advance in intellectual history, with broad implications for humanists and social scientists. But it is also, of course, a complicated and controversial topic, often fused with concerns about the growing power of Silicon Valley. To avoid complicating an already complex argument, I have deliberately limited this article to older quantitative methods (scatterplots, logarithms, curves). None of those things were invented in the Bay Area. The charged question they raise is not about computers but about the possibility of dialogue between humanists and social scientists. I suspect these are also the real stakes of the larger debate about digital humanities, but where debates are this impassioned and complex, it is wise to proceed one step at a time.

So the main technological innovation discussed in this article will be a very old one: Arabic numerals. We needed numbers to trace changes in the pacing of fiction, because those changes only became legible after averaging many passages. Numbers can extract a trend from noisy variation. On the other hand, because this is a powerful technique, it makes differences visible everywhere: in the real world, no group of books has a mean pace exactly the same as another. It thus becomes

necessary to evaluate evidence cautiously, comparing the trend we have identified to the magnitude of random variance. Social scientists have developed standard ways of doing this; one widely-used measure is the Pearson correlation coefficient, or r . Applying that equation to the pattern in Figure 2, we discover that the correlation between a volume's date of publication and the mean of $\log(\text{passage duration})$ for that volume is $-.64$. The negative sign indicates that pace goes down as the date goes up. By loose scientific convention, absolute values of r around $.1$ are called small effects; those around $.3$ are medium-sized effects; those greater than $.5$ are large effects.²⁰ We can also evaluate the statistical significance of the correlation, which is about eight orders of magnitude below the $.05$ threshold commonly used as a maximum. In reality, statistical significance is rarely a sufficient test for distant reading, because if there is enough evidence, almost every pattern becomes *statistically* significant. Effect size is a more relevant question, and this is a large effect.

But the real tests of a quantitative argument are never contained in a single number anyway. More often, we want to know whether the pattern holds up under different assumptions, and viewed from different angles. One advantage of computing is that it becomes easy to pose these skeptical questions by subdividing the evidence and comparing different perspectives. The code and data underlying this article are available online for researchers who want to do that.²¹ We might, for instance, test alternate metrics. The pattern in Figure 2 holds up if we take the median for each volume, rather than the mean. Or we might try alternate methods of sampling. In this experiment, we always included the first two and last two passages in each book. So the beginnings and ends of each story got read much more often than they would have if we were sampling entirely randomly. That's one way of modeling pace, and you might like that model, if you think the beginnings and ends of stories are important. But we could alternatively down-weight those passages to reflect the fact that they're sampled from a small portion of the text. Figure 3 shows the pattern we get with that method. The trend is still clear, and still basically the same, but a little less dramatic: it becomes a 70-fold decline instead of a 100-fold ($r = -.53$, $p < .000001$). I think sampling evenly across a whole narrative is easier to explain, so this is the model I have used in figures from here on. You will also notice that I have added a gray area around the trend line to reflect uncertainty about the real location of the mean.

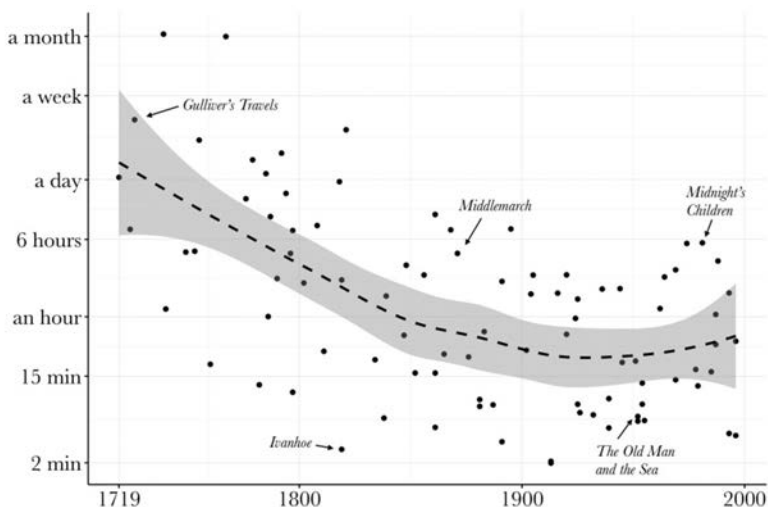


Figure 3. The average length of time described in 250 words of narration. The average is still calculated as $\text{mean}(\log(\text{passage_duration}))$, but the first two and last two passages are now down-weighted to reflect over-sampling of the beginning and end of stories. The shaded ribbon represents a 95% confidence interval for the dashed curve, which is itself calculated by loess regression.

Another set of alternate perspectives we might want to test involves the set of books used in the experiment. These questions often get especially close scrutiny, because Moretti's turn-of-the-century articles presented distant reading as a way of getting beyond the literary canon—leading to a long argument with book historians, who had their own tradition of getting beyond the canon by precisely mapping the circulation of editions.²² Although it may sound plausible that researchers should resolve these debates and agree on a representative sample of texts before proceeding, I think that goal can be a red herring. The literary past can be sampled in many ways; no single sample is appropriate for all questions. We should compare different samples, and computation makes it easy to do that by subdividing the evidence. But for many questions—like this one—the contrast between canonical works and obscure ones isn't the point, and doesn't make an enormous difference. Long diachronic patterns tend to be robust.²³ Narrative pace will vary subtly across different parts of the literary field, and those variations are worth exploring in a later phase of research. But when literary practice as a whole moves broadly in the

same direction, it is not always necessary to minutely map synchronic variation before acknowledging the diachronic trend.

A final skeptical question we might pose involves the alternate perspectives of the subjects in the experiment itself. The passages aggregated above were read, after all, by three different people, with different perceptions of time. So the three of us also read six novels in common and compared our reactions to the same passages. We found a lot of agreement; across the whole set of shared passages, paired estimates of time correlated with each other at $r = .78$. But this also leaves room for significant disagreement, ranging from slightly different guesses about the time it takes to eat dinner, to puzzles about the beginning of *Ivanhoe* (1820), where Walter Scott casually juxtaposes a twelfth-century forest to its modern remnants (which may or may not be part of the story).

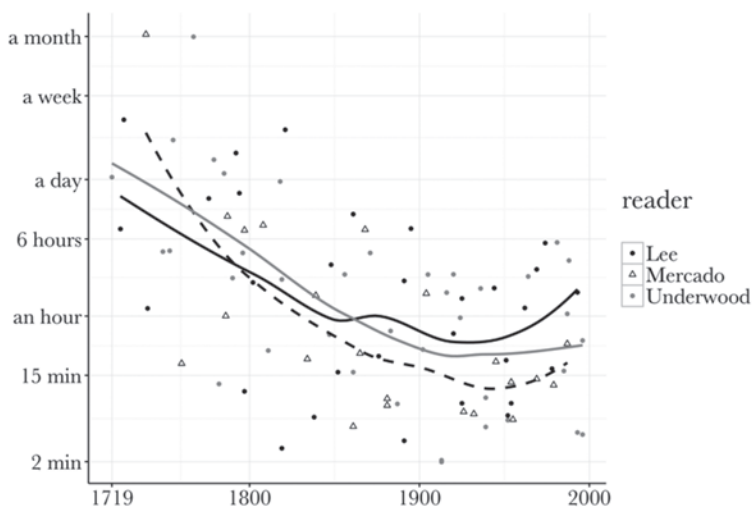


Figure 4. The perspectives of three different readers. Scale otherwise as in figure 3. The means for Lee, Mercado, and Underwood are represented by black, dashed, and gray lines, respectively.

Does Scott's forest stretch across 600 years of narration? There is no right answer to that, or to most of the other questions we confronted. The representation of time in fiction can be dizzyingly subjective, and its deviations from linear order have long interested narratologists.

These slippery questions are absolutely an important topic for research. But it is also true that three different readers mostly agreed with each other about the duration of many different passages. And when we backed up to average all the passages in each book, and plot a curve across historical time, the three of us almost entirely agreed about its trajectory ($r = .96-.97$). This area of intersubjective agreement is also an important topic for research, and one that literary scholars have not yet explored.

I want to pause to underline the word intersubjective. We inherit an assumption that quantitative methods produce objectivity, or at least pretend to. The history of that straw man is matter for another day, but let me briefly offer my view—which is that numbers are no more objective than words. They are just signs that allow human observers to wrestle with questions of degree. As we back up and look at large historical patterns, we need numbers, not because we are trying to be more objective than usual, but because our material is becoming more varied and comparisons are becoming difficult. We need a way to acknowledge variation, uncertainty, and intersubjective disagreement. These complications, often adduced as a reason why numbers could never describe literature, are exactly why we need numbers to understand long literary timelines.

III: INTERPRETING A MODEL

After testing this pattern from many different angles, we now have a fairly robust model of the relation between historical time and narrative pace. Our model is simply the curve in Figure 3. Like other statistical models, a curve is a simplified description of a relation between variables. This curve may look slightly different from different angles, but it doesn't dissolve. So how should we interpret it?

I'm writing "interpret," not "explain," because I don't think causal questions are always the most interesting ones to ask at this scale. A three-century trend spanning the Atlantic Ocean is likely to have many causes. It will also be hard to separate cause from effect, since the feedback loop from cause to effect to cause again will have cycled through many times in the course of 300 years. We have been trained to insist, a bit sententiously, on the difference between correlation and causation. That difference matters enormously if you're testing the effect of a new drug. But when you're studying a self-reinforcing cycle across three centuries, it can be close to moot.

So I will discuss causality only briefly, and mainly to fend off a couple of tempting assumptions. The first of these is Genette's notion that narrative pace changes when writers just drop the summary and replace it with ellipsis. I don't think there is much truth to that hypothesis. For one thing, the division between scene and summary is never crisp. Looking at the distribution of passages across different lengths of time (Figure 5), I wouldn't know where to divide scene from summary. In the eighteenth century, there seem to be three distinct humps in the distribution (the middle one corresponding roughly to the concept of a week).

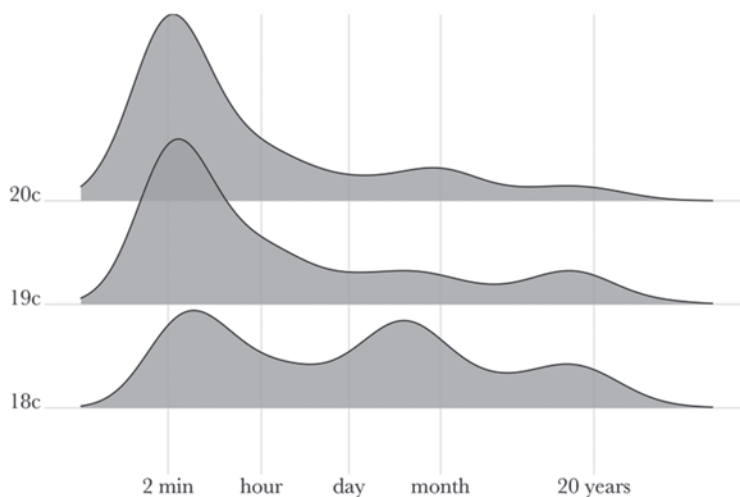


Figure 5. Time narrated in 250-word passages. A density curve is estimated for individual passages (not whole books) in each century.

It is true that ellipses become more common in the twentieth century. Genette is right about that. But outright gaps in narration remain a rare phenomenon—too rare to explain the broader change of pace. In fact, the books that lead the way by introducing a slower pace are not modernist novels experimenting with ellipses, but, if anything, nineteenth-century bestsellers like *Ivanhoe* and *East Lynne* (1861).

The second tempting causal explanation I need to fend off may be familiar from E. P. Thompson. Industrial capitalism certainly changes

the organization of time in the process of disciplining work. Thompson famously connects this to the diffusion of watches, which help people measure smaller units of time. It seems intuitively plausible that fiction might have adapted itself to this more finely-divided, closely-regulated, industrial world.²⁴

And indeed that may be some part of the explanation for temporal compression.²⁵ But it can't be the whole explanation, because the trend we see in fiction isn't found in other forms of narrative. If this were just a broadly social change in the experience of time, we might expect it to affect nonfiction as well. For instance, biographies and autobiographies are also narratives centrally concerned, like novels, with individual experience. But in those genres, the pace of narration doesn't seem to change as it changes in the novel. I haven't tagged as many biographies, and they are very diverse, so I can't say confidently whether the pace of narration increases or decreases in biography. That is why the shaded band in Figure 6 is so wide for that genre. But certainly there is no evidence yet for the kind of clear deceleration we see in fiction.

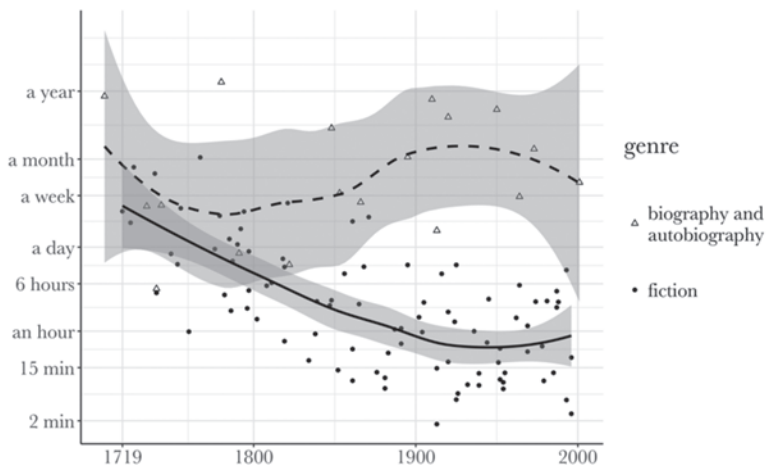


Figure 6. Weighted average of 250-word passages in fiction and biography.

In fact the changes we see in fiction might be caused by what Emile Durkheim called structural differentiation—which leads institutions to specialize in distinct social niches.²⁶ In other research, I have encountered further evidence that biography and fiction were differentiating; I won't go into it deeply here because I have promised not to rely on machine learning in this article, but let's just say that as we proceed down the timeline, predictive algorithms find it easier to tell biographies and works of fiction apart in a blind taste test. The genres become textually less similar in many ways, not just in their relation to time.

So, I admit, it is tempting to explain changes in narrative pace as a symptom of a general differentiating process that separated fiction ever more firmly from other genres of narrative. But I will resist the temptation to advance that as a causal argument, at least for now. We haven't fully described this trend yet. Instead of rushing to explain it causally, let's thicken our description by looking for textual patterns and literary pleasures bound up with the change. That can be another way of answering the *why* question. It may not explain agency—Aristotle's efficient cause—but it could help us understand purpose, or formal cause. What, in other words, did writers achieve by compressing fictional time?

One pleasure bound up with this change is overt and naïve. I'm going to call it breathless narration—a narrative strategy that draws attention to the compression of time and presents it as a source of thrilling surprise. Perhaps the ur-text for this is Samuel Richardson's *Pamela* (1740), where narration can be literally breathless because its epistolary context is dramatized: I have to finish writing before Mr—bless me, there is Mr B at my door right now! But those tricks have limits; eventually you get mocked by Henry Fielding. By the late eighteenth century, novelists were working out a subtler and more durable stance of breathlessness. Maybe narrators can't literally describe events as they happen, but they can still describe events in a way that highlights their immediacy. This is one of the things that makes Frances Burney sound more modern than Daniel Defoe; she keeps reminding her readers that events are succeeding each other quickly. A new paragraph will begin “A few moments after he was gone[.]” or “At that moment Sir Robert himself burst into the Room[.]”²⁷ Ann Radcliffe similarly intensifies events by making them happen “at the very moment” something else does.²⁸

The gesture feels naïve because it attempts something that narrative cannot literally achieve. Writers have many ways to create intensity—diction, suspense, conflict. But breathless narration seems to create

intensity by impressing the reader with the number of events crammed into a small space of represented time. The problem is that we don't directly experience represented time, and aren't necessarily startled by its compression. The editor of a music video can overwhelm a viewer with rapid jump cuts. But if the eventfulness of narration really began to startle a reader, they could just turn pages more slowly. And yet, breathlessness is a very popular narrative stance—so popular that on some level it must work.

Although breathlessness may seem naïve, it is often flagged self-consciously—for instance when characters do something “in much less time than it takes to write it.”²⁹ Acknowledging the difference between space on the page and fictional time may sound like an apology. But more often, the breathless narrator is boasting about the brevity of represented time. You can feel this sort of pride when Philip Marlowe writes, at the end of *The Big Sleep*: “I rang the bell. It had been five days since I rang it for the first time. It felt like a year.”³⁰ Here Marlowe anticipates the reader's sense that many narrative twists have been packed (a bit implausibly?) into a short space of represented time, and turns that compression into yet another occasion for hard-boiled weariness. L. A. is a tough town for a narrator; you really earn that 50 dollars a day plus expenses.

But while the self-consciousness of breathless narration is odd and interesting, it probably isn't the main pleasure produced by the steadily slowing pace of fiction over the last three centuries. The compression of time also accompanied bigger, simpler shifts of emphasis. In particular, it was strongly associated with dialogue and with insistently physical description.

The effect of dialogue is easy to understand: since it takes roughly the same time to read dialogue that it would take to speak it, dialogue necessarily slows down the passage of represented time.³¹ Since this effect is fairly obvious, I thought the slowing pace of fiction would correlate very closely with the amount of space spent on dialogue, and there is in fact a medium-sized correlation ($r = -.28$).

But changes in narrative pace correlate even more closely with rising emphasis on concrete description. There are many ways to measure this. For convenience, I'm going to borrow a measure that was developed in the Stanford Literary Lab. It is simply a list of words that tend to be used in physical description: directional prepositions, like “up” or “down,” physical verbs like “walk,” physical adjectives, numbers, and body parts.³² A list like this isn't a perfect or stable measure. Words have multiple meanings, and the definition of concreteness can change

across time. But the changes aren't huge, and as Figure 7 shows, this list works in practice very well across three centuries, organizing biographies and fiction as a single pattern ($r = -.60$).³³ This tendency for slow narration to accompany sensory, spatial description may be why we assume that a critical method organized around brief spans of time will go hand in hand with concrete particularity. That pairing isn't inevitable: in principle, you could spend a page slowly describing ten seconds of abstract thought, and some novelists do. But usually, slowing the pace of narration to a scale of minutes has meant increasing the amount of sensory detail.

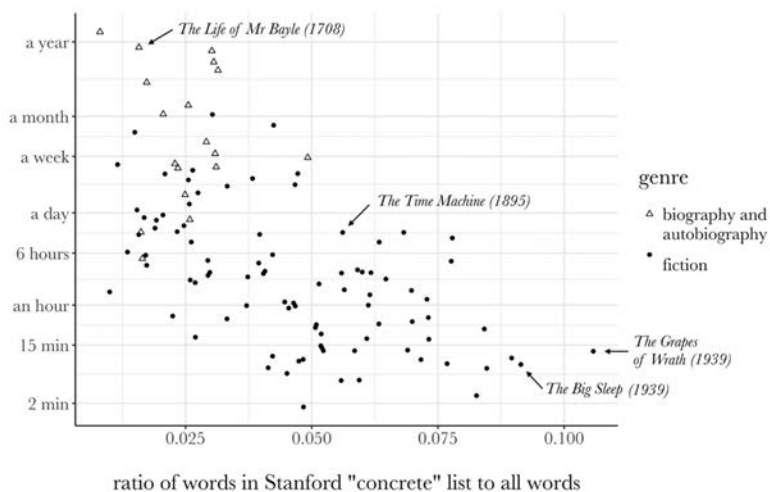


Figure 7. The pace of narration, and the prevalence of concrete diction. Each point represents a volume of biography or fiction. A slower pace correlates with more concrete description ($r = -.60$). Both things are also more common in twentieth-century fiction than in biography, or earlier fiction.

This tells us, perhaps not *why*, but *how* the pace of fiction changed, and what pleasures were thereby produced. The slowing pace of fiction is strongly bound up with the prominence of dialogue and physical description, across a span of 300 years. I have described this shift as an attempt to emulate the immediacy of dramatic presentation. That's not a radically new idea; in the 1920s, Percy Lubbock already saw that fiction had been moving toward "showing" rather than "telling" in the

late nineteenth century.³⁴ But now, we have stepped far enough back to see Lubbock's argument as part of a much bigger picture. This was not, as Lubbock suggests, an innovation made by Gustave Flaubert or Henry James; it was a 200-year trend that transformed time in fiction, moving it steadily from a scale of days to a scale of minutes. Popular fiction, like Scott's *Ivanhoe* and Ellen Wood's *East Lynne*, played as big a role in this process as James did.

IV: THE TIMELINE AND THE ANECDOTE

The New Historical anecdote often runs longer than 30 minutes. It is not literally the same thing as a fictional scene. Nor did it rise to prominence in literary criticism at exactly the same time as the (eighteenth- and nineteenth-century) shifts in narrative we have been tracing. But I do think the preference for brief scenes in fiction eventually produced a similar preference in literary criticism. Changes in the pacing of fiction accompanied, and helped consolidate, a belief that literature is distinguished from other forms of language by attention to the immediacy of individual experience. By the middle of the twentieth century this insistence on immediacy was also defining a regulative boundary between literary criticism and other academic disciplines.

The opening pages of this article have already sketched a few links in this genealogy. Auerbach, for instance, explicitly based the structure of *Mimesis* on modernist writers' treatment of time. Gallagher and Greenblatt similarly invoked the temporal condensation of the realist detail to justify their own anecdotal method. The New Critics don't always insist overtly on temporal compression (in part because you hardly need to belabor that point if you're writing about lyric poetry), but they too defined literature through its qualities of immediacy and concreteness.

All of these critical texts are at least a few decades old. But the notion that literature is defined by temporal compression has lost none of its power: it still governs critical discourse today. A good example can be found in the introduction to a recent special issue of *Genre*, where Jesse Rosenthal translates controversies about the scale of distant reading into a deeper struggle between "narrative" and "data."³⁵ For Rosenthal, the power of data is displayed in visualization, which he understands as essentially timeless. Dr. John Snow's famous map of cholera, he points out, "works by collapsing the temporal dimension," so we can see all the cholera cases in September 1854 as points in a single image.³⁶ By contrast, "the experience of narrative is one that takes

place over time,” so “narrative will tend to resist this sort of significance-through-aggregation.”³⁷ A novel inevitably treats different illnesses as separate moments in a sequence: temporality leads to individuation.

One could quibble about the implicit assumption here that images are perceived all at once. In practice, timelines are read from left to right, rather like a line of type. But Rosenthal is introducing a special issue about the Victorian novel, and he is basically right about differences of scale in the nineteenth-century examples he considers. John Snow’s map does compress a whole month into a single image. By contrast, Victorian novels unfold on a scale where each telltale cough, each fever, each touching episode of delirium must be described separately.

But I hope this essay has prepared readers to see that the Victorian novel’s particular approach to pacing is not the same as that of narrative or literature in general. Consider, for instance, Defoe’s *Journal of the Plague Year*: “[T]he next week there seemed to be some hopes again; the bills were low, the number of the dead in all was but 388, there was none of the plague, and but four of the spotted-fever.”³⁸ For Defoe’s narrator, a week of contagion killing 388 people can be a narrative unit just as reasonable as a single pathetic cough might be for Charles Dickens. On some pages, the narrator backs up to give us a table that covers actuarial figures for several parishes across a whole month—a scale of aggregation closely comparable to John Snow’s street map of London.

A Journal of the Plague Year is not exactly a typical work of fiction, but it is not an outlier either. Actuarial tables may not be common in novels, but the evidence surveyed above shows that it is quite common for early-eighteenth-century fiction to treat weeks and months as narrative units. In light of that evidence, it should be clear that there is nothing timeless, hostile to narrative, or inherently unliterary about a map of London that covers all of September 1854. September is just a unit of time a bit larger than novelists and literary critics have recently taken as normative.

In placing Rosenthal next to Auerbach, Gallagher, and Greenblatt, I am taking him as characteristic of literary study at its best. Although Rosenthal’s introduction is titled “Narrative Against Data,” it is never one-sided; in fact, it characterizes the encounter of “narrative” and “data” as “an exciting one.”³⁹ But at the same time, it is centrally concerned to define “data-driven approaches” as “a different thing altogether” from “traditional literary criticism.”⁴⁰ I think this ambivalent response to distant reading typifies contemporary literary study. We

are an open-minded discipline, but also nervous about the porosity of our borders. Many literary scholars are intrigued by quantitative methods, and willing to see them explored—as long as a protective firewall of some kind can be guaranteed. For instance, we could posit in advance that numbers are incapable of capturing the truly literary aspects of literature, such as fine divisions of time. If numbers are about months or years, and narratives about minutes, then it might be safe for literary scholars to work with data, precisely because data can never touch the beating heart of our subject.

The impulse to reinforce disciplinary borders is by no means unique to Rosenthal. Many other scholars have emphasized the brevity of literary experience in order to distinguish properly literary analysis from social science and historical narrative. Gallagher and Greenblatt, for instance, described New Historicism as a systematic reassertion of the brief Auerbachian fragment against Marxist and Annaliste generalization. Perhaps taking a long view of these controversies can reduce some of the tension surrounding them. Disciplines change slowly, and the real stakes of theoretical debate are often lower than we pretend. Although quantitative social science has loomed on the horizon of the humanities for many decades, large scales of quantitative analysis have never replaced other approaches to the human past, and the inroads they are making in the humanities today remain rather marginal and modest.

Moreover, the pleasures that currently attract students to literary study guarantee that the literary humanities, in particular, will remain centered on a phenomenological scale of description. This article has criticized presentist assumptions about the scale of literary experience. Literary time was not always measured in minutes; narrators have not always insisted so fiercely on the concrete particularity of brief events. But characterizing these assumptions as presentist cuts both ways, since the present is after all where we spend most of our lives. In the process of showing that literature was not always defined by temporal immediacy, I have also acknowledged that temporal immediacy is now an important boundary separating literary language from nonfiction. Students become English majors, and eventually English professors, because they love vivid fragments of individual experience. So distant readers who want to persuade English professors will need to keep imitating the strengths of post-Romantic literature. We will have to keep using anecdotes and close readings, as well as graphs and maps, when we address a disciplinary audience.

In short, there is very little danger that recent quantitative experiments will displace a conception of literature that has been three centuries in the making. Literary scholars' preference for close description of finely-divided moments is already buttressed by the recent history of literature itself. It doesn't need to be further buttressed by universalizing arguments about the proper scale of truly literary analysis. The real dangers for our discipline lie elsewhere, and erecting higher walls between "narrative" and "data" will not protect us from them. The future that should worry literary scholars is not that quantitative methods will spread too rapidly in our discipline. It is more likely that the inertia of literary curricula will repel new questions requiring statistical training only too effectively. In that case, exciting discoveries about the *longue durée* of literary history could be made in other disciplines—like Communications or Information Science—where quantitative research finds a more receptive curricular soil.

That is a likely future, and not necessarily a bad one, but I don't think it is the best one we could create. I would rather see distant reading find at least part of its home in literary studies. But that can only happen where literature departments are confident enough to stop building firewalls against quantitative social science, and instead assimilate numbers as one part of a diverse toolkit. In any case, I plan to continue working across this divide. I see close readings and statistical models not as competing epistemologies but as interlocking modes of interpretation that excel at different scales of analysis. Histories of broad trends will often need to pause for close description—for instance, of a passage or two from Burney—in order to understand the human significance of a trend. But the converse is also true. If we want to understand how case studies fit together to form literary history, we will need more than our unaided memories. Even scholars who are interested strictly in canonical writers will sometimes need to measure questions of degree across a long timeline. Otherwise we can end up giving Joyce and Proust credit for changes that actually sprawled across several centuries.

Far from threatening each other, quantitative and qualitative interpretation are mutually illuminating. In fact, this article has tried to show that large-scale quantitative analysis can cast light even on our reasons for valuing brief scenes and gem-like moments. The timeline and the anecdote are not just compatible, but complementary ways of thinking about the past; each perspective needs the other to better understand itself.

University of Illinois, Urbana-Champaign

Ted Underwood

NOTES

This article was sparked by thinking about Sharon Marcus, "Erich Auerbach's *Mimesis* and the Value of Scale," *MLQ* 77.3 (2016): 297–319. Eleanor Courtemanche contributed many leads on narrative theory; Andrew Goldstone convinced me to try sociological content analysis. Texts of novels after 1900 came from Hoyt Long and Richard Jean So at the Chicago Text Lab; for texts before 1900, I collaborated with HathiTrust Research Center. Data about the novels was gathered in collaboration with Jessica Mercado and Sabrina Lee; research was funded by SSHRC via the NovelTM project, directed by Andrew Piper.

¹ See Catherine Gallagher, "Formalism and Time," *MLQ* 61.1 (2000): 229–51.

² Walter Pater, *Studies in the History of the Renaissance* (London: Macmillan, 1873), 210.

³ Gallagher and Stephen Greenblatt, *Practicing New Historicism* (Chicago: Univ. of Chicago Press, 2000), 26.

⁴ Erich Auerbach, *Mimesis: The Representation of Reality in Western Literature* (New York: Doubleday, 1957), 484.

⁵ Gallagher and Greenblatt, 42, 51.

⁶ See Sandra M. Gilbert and Susan Gubar, *The Madwoman in the Attic: The Woman Writer and the Nineteenth-Century Literary Imagination* (New Haven: Yale Univ. Press, 1987); Geoffrey Hartman, "Thesis: The Halted Traveler," in *Wordsworth's Poetry, 1787–1814* (Cambridge: Harvard Univ. Press, 1971), 1–30; and Greenblatt, *Learning to Curse: Essays in Early Modern Culture* (New York: Routledge, 1990).

⁷ For more on this leap across scales, see Marcus.

⁸ Hartman, 12. See James Chandler, *England in 1819: The Politics of Literary Culture and the Case of Romantic Historicism* (Chicago: Univ. of Chicago Press, 1998).

⁹ Edgar Allan Poe, "The Poetic Principle," *Home Journal* 36 (1850): 1.

¹⁰ See Auerbach, 483–84.

¹¹ David Harvey influentially argues that modernism "is . . . suffused with the sense of the fleeting, the ephemeral, the fragmentary, and the contingent" (*The Condition of Postmodernity: An Enquiry into the Origins of Cultural Change* [Cambridge: Blackwell, 1990], 11). See also Erin D. Sells, *The Story of a Day in Modernism* (PhD diss., Emory Univ., 2009), 9, 135–36.

¹² Gérard Genette, *Narrative Discourse: An Essay in Method*, trans. Jane E. Lewin (Ithaca: Cornell Univ. Press, 1980), 95.

¹³ See Auerbach, 483; and Genette, 97–98.

¹⁴ Genette, 106.

¹⁵ See Seymour Chatman, *Story and Discourse: Narrative Structure in Fiction and Film* (Ithaca: Cornell Univ. Press, 1989), 71–78.

¹⁶ "[I]t is obvious that summary remained, up to the end of the nineteenth century, the most usual transition between two scenes, the 'background' against which scenes stand out, and thus the connective tissue par excellence of novelistic narrative, whose fundamental rhythm is defined by the alternation of summary and scene" (Genette, 97).

¹⁷ Distant reading is commonly presented as a mode of interpretation that would be needed only for texts and archives that are unreadable through obscurity, complexity, or sheer scale. Natalia Cecire argues that it "depends on the premise that [the works it considers] are generic and stylistically unremarkable" ("Ways of Not Reading Gertrude Stein," *ELH* 82.1 [2015]: 282). I am suggesting that distant reading addresses the limits of our capacity, not to read, but to remember and compare. It can help us understand the relations between works, even when those works are individually quite remarkable.

¹⁸ See Janice Radway, *Reading the Romance: Women, Patriarchy, and Popular Literature* (Chapel Hill: Univ. of North Carolina Press, 1984).

¹⁹ For a longer discussion of distant reading as an experimental method, see Ted Underwood, "A Genealogy of Distant Reading," *Digital Humanities Quarterly* 11.2 (2017): <http://www.digitalhumanities.org/dhq/vol/11/2/000317/000317.html>.

²⁰ See Jacob Cohen, "A Power Primer," *Psychological Bulletin* 112.1 (1992): 155–59.

²¹ See Underwood, "tedunderwood/moments: Code as used in ELH article. (Version v1.0)," Zenodo, <http://doi.org/10.5281/zenodo.1137728>.

²² An eloquent argument for bibliographically stable samples can be found in Katherine Bode, "The Equivalence of 'Close' and 'Distant' Reading; or, Toward a New Object for Data-Rich Literary History," *MLQ* 78.1 (2017): 77–106.

²³ For another example of a long, robust trend, see figure 1.2 in Mark Algee-Hewitt, Sarah Allison, Marissa Gemma, Ryan Heuser, Franco Moretti, and Hannah Walser, "Canon/Archive: Large-Scale Dynamics in the Literary Field," *Stanford Literary Lab Pamphlets* (2016), <https://litlab.stanford.edu/LiteraryLabPamphlet11.pdf>.

²⁴ See E. P. Thompson, "Time, Work-Discipline, and Industrial Capitalism," *Past & Present* 38 (1967): 56–97.

²⁵ For a more fully worked-out account of the connection, see Sue Zemka, *Time and the Moment in Victorian Literature and Society* (Cambridge: Cambridge Univ. Press, 2012).

²⁶ See Emile Durkheim, *The Division of Labor in Society* (New York: Free Press, 1964). See also Dietrich Rueschemeyer, "Structural Differentiation, Efficiency, and Power," *American Journal of Sociology* 83.1 (1977): 1–25.

²⁷ Frances Burney, *Cecilia; or, Memoirs of an Heiress*, 2 vol. (London: George Bell, 1890), 1:357, 358.

²⁸ Ann Radcliffe, *The Mysteries of Udolpho* (Oxford: Oxford Univ. Press, 1992), 153.

²⁹ William M. Baker, "Mr Dod's Six Shots," *Harper's New Monthly Magazine* 32 (1866): 210.

³⁰ Raymond Chandler, *The Big Sleep* (London: Penguin, 2011), 227.

³¹ For more insight on the growing role of dialogue in fiction, see Marissa Gemma, Frédéric Glorieux, and Jean-Gabriel Ganascia, "Operationalizing the Colloquial Style: Repetition in 19th-Century American Fiction," *Digital Scholarship in the Humanities* (2015): <http://dx.doi.org/10.1093/llc/fqv066>; and Grace Muzny, Algee-Hewitt, and Dan Jurafsky, "Dialogism in the Novel: A Computational Model of the Dialogic Nature of Narration and Quotations," *Digital Scholarship in the Humanities* (2017): <https://doi.org/10.1093/llc/fqx031>.

³² Ryan Heuser and Long Le-Khac, "A Quantitative History of 2,958 Nineteenth-Century British Novels: The Semantic Cohort Method," *Stanford Literary Lab Pamphlet Series*, <https://litlab.stanford.edu/LiteraryLabPamphlet4.pdf>.

³³ Heuser and Le-Khac have already shown that concreteness increases steadily in nineteenth-century fiction, so one might at first wonder whether concreteness is simply a proxy for date. But a variable that was simply a proxy for date wouldn't also distinguish biographies from fiction.

³⁴ Percy Lubbock, *The Craft of Fiction* (New York: Viking Press, 1957), 62, 74. Heuser and Le-Khac also explain the trends they observe as a move from telling toward showing.

³⁵ See Jesse Rosenthal, "Introduction: 'Narrative Against Data,'" *Genre* 50.1 (2017): 1–18.

³⁶ Rosenthal, 11.

³⁷ Rosenthal, 11.

³⁸ Daniel Defoe, *A Journal of the Plague Year* (London: J. M. Dent, 1957), 5.

³⁹ Rosenthal, 16.

⁴⁰ Rosenthal, 9.