

# Shifting Scales: Between Literature and Social Science

James F. English and Ted Underwood

---

**S**cale, from the Latin *scala* (ladder) via Old French *escalier* (climb), refers to valuation (“a graduated range of values forming a standard system for measuring or grading”) as well as to dimensionality (“the relative size or extent of something”) (*Concise Oxford English Dictionary*, 12th ed., 1282). We have assembled this special issue of *MLQ* to highlight important new work of literary-historical inquiry, partly though not exclusively pertaining to the digital humanities, in which problems arising from the nexus of scale and value have become conspicuous concerns of method. In doing so, we are not implying that our discipline’s current perturbation over these problems is without precedent. Every field of study must, at every stage in its development, define its proper scope, locating workable boundaries between its own objects, zones, and tools of research and those external to it. To engage in literary studies has always meant operating in a specially constructed and privileged space of the “literary,” a field of practice scaled to incorporate everything relevant to an understanding of literature, while redlining adjacent neighborhoods as precincts of extra- or subliterary concern, subject to the authority of other disciplines. Within this discipline-shaped space, we rely on a host of other scalar constructs. Certain temporal spans (century, literary period, artistic generation), geocultural categories (national literature, regional literature, diasporic or exilic literature), formal entities (protagonist, genre, individual work), and so forth supply us with the basic units we need to organize our research projects and structure our intellectual and institutional divisions of labor. The degree of prestige attaching to these built units of study is hierarchized, such that some national literatures,

*Modern Language Quarterly* 77:3 (September 2016)

DOI 10.1215/00267929-3570612 © 2016 by University of Washington

historical periods, genres, and individual works hold more value on the field than others; they carry more weight and exert greater influence than the units around them. And while the bulk of normative scholarship operates comfortably in accord with prevailing scales and values of scholarly practice, much of the excitement has arisen from the impulse to challenge them.

Indeed, a familiar narrative of our discipline's history presents it as essentially a drama of competing scales. The founding schools in the early decades of the last century all sought to establish literary studies as a distinct and legitimate—teachable, testable, rigorous—field of higher learning via radical contractions of the set of objects relevant to the enterprise. Different though they were, the movements that emerged in Moscow, Cambridge, and Baton Rouge all fought for much narrower specifications of the literary object than had been on offer previously through philology or *belles lettres*. They distinguished it sharply from other kinds of linguistic and cultural artifact, isolated it from the social contexts of authorship and publishing, largely withdrew it from the history of languages, and all but severed its connection to ancient models. Along with this reduction to the range of relevant objects came new methods privileging smaller units of scholarly analysis: most famously the isolate “poem itself” (Brooks and Warren 1976 [1938]: 13), enshrined by the New Critics, but also subsidiary units meant to capture the truly “poetic” elements of a text, such as the “device of enstrangement” for the Russian formalists (Shklovsky 1990: 9), and companion units such as the brief “protocol” of individual response, whose assessment alongside the poetic text shaped the terms of practical criticism (Richards 1929). Scalar contraction was, for half a century, a winning strategy. Literary studies massively expanded its institutional footprint and widened its cultural power while for the most part maintaining an orthodoxy of methodological constraint with respect to the scope of its analytic procedures and its expertise.

Then, in the final quarter of the century, the great contraction ended, producing a kind of big bang effect that accelerated the dilation of the analytic aperture of literary studies. Having isolated the “poem” as its *raison d'être*, justifying its place in the academy on the basis of the specialized forms of expertise required to understand the inner workings of literary art, the discipline began to admit for scholarly study the

very objects against which the poem was defined: genre fiction and comic books, popular-music lyrics, products of global screen culture or of the fashion and advertising industries. Having pushed authorial biography outside the compass of concern, the discipline launched major subfields based precisely on the social identities of authors: women's literature, African American literature, gay and lesbian literature. Having all but killed off the discipline of philology, literary studies claimed much of that discipline's old territory for itself, reviving attention to the history of books and manuscripts and textual variants, to systems of print and publication, institutions of sale and circulation, practices of reading and reception. Having scrupulously sequestered literary study from the study of extra-literary history, the discipline began to reorganize much of its research agenda under the motto "Always historicize!" (Jameson 1981: 9), a turn that eventually led scholars to incorporate into their literary research texts and artifacts as seemingly remote as medical case studies, juridical documents, ships' logs, field notes, and shopping lists. By the end of the century, what had begun as a project of narrow formal specification, insistently modest in scope, had burgeoned into the most capacious of disciplines, scaled to support research into archives so vast and heterogeneous as to seem coextensive with the territory of the humanities themselves.

This is admittedly a cartoonish history, treating as absences what were in fact vital but less valued countertendencies and failing more generally to place enough weight on the value side of the scale-value equation. The drive toward a wider program of literary research was always also a struggle over regimes of cultural status and standards of institutional access, and we know that in many respects our discipline remains even today a narrow and exclusionary one. Still, most of us would probably concede the thrust of the narrative, that the field has been expanding conceptually and analytically, widening its purview and increasing exponentially its stock of raw material, for much of the last half century.

One way to view the discipline's current intense engagement with problems of scale is as an apotheosis of this expansion. Indeed, Eric Hayot (2013) says that the discipline is facing a "*crisis of largeness*." He observes that, since the turn of the twenty-first century, "all (or almost all)" of the major developments in literary studies have sought "to substantially *enlarge* the scale at which we discuss literature and literariness."

His account of these developments suggests why their impact might be more disruptive than what occurred at earlier moments of disciplinary build-out: while previous expansions widened the field that we collectively produce, they did not force us individually to work with larger units of analysis, certainly not with the immense units that are now in play.

To be sure, an expansion or contraction along one axis of research generally gives rise to new scales, standards, and units of measure along other axes. The broadening of the discipline space of literary studies from the 1970s through the 1990s did bring some new, larger units to the forefront of research, as when early attempts to understand the cultural role of television occasioned a shift from the individual program as the primary unit of analysis to what Raymond Williams (1989 [1971]: 133) termed the “general flow” of programming: “the organization, the methods and the values within and through which particular programs occur.” If widely adopted, this shift would have discouraged literary scholars from simply substituting a single television show for a novel or play so as to produce an article-size “close reading” of a soap or sitcom. But the continued dominance of this latter practice indicated just how deeply ingrained was the methodological preference for close scrutiny of text-size units. New Historicism, to take another example, extended literary-historical research into a vast new set of archives but also (notoriously) privileged very small units of textual and temporal analysis. Its essential method was “what in optics is called foveation, the ability to keep . . . a tiny, textualized piece of social behavior within the high-resolution area of focus” (Greenblatt 1997: 18) even while pursuing a grander line of inquiry. This was true of New Historicism’s handling of temporal as well as textual units, as when it endeavored to grasp the discursive logic of an entire literary period by making a synchronic deep dive into “a single year’s writings” (Chandler 1998: 32). For all its expansive effect on the texts and topics deemed pertinent to literary studies, New Historicism was in this respect a “nanohistoricism” (Liu 2008: 5).

Today, by contrast, we confront something more like gigahistoricisms. The system of discrete and contrasting literary periods, among the discipline’s most durable devices (Underwood 2013), has been put under pressure by long-span approaches such as Rita Felski’s (2008; 2011: 574) study of the “transtemporal” dimensions of literary response,

Laura Doyle's (2015: 338) "longer *durée* dialectics" for postcolonial literary studies, and Wai Chee Dimock's (2006: 225) probing of American literature's "deep time" backstory, a multicentury project that gestures toward the still more extreme temporal units of "geological and cosmological time," deployed in various fashion by practitioners of evolutionary criticism (Boyd, Carroll, and Gottschall 2010) and the ecocritical mode of environmental humanities (Buell 2007; Heise 2014). With their invocations of the "planetary," some of these deep-time versions of literary history have contributed as well to a decisive enlargement of geographic units. The division of the field into relatively discrete national literatures, still common practice in the 1990s, has begun to seem downright archaic in light of reorientations toward transnational and transoceanic research and the general effort, highlighted by David Damrosch (2011: 20), to cultivate a "fully global perspective" on literary production and exchange. This effort has of course gained substantial impetus from the work of Franco Moretti (2013: 49, 105), whose provocative conjectures on world literature and its history entail analysis of formal units "much larger" than the individual text, as large even as the "world literary system" itself, or what Pascale Casanova (2005: 12), in her own influential approach to literature's global history, calls the world literary "space" or "structure," and what Alexander Beecroft (2015: 17–21) further elaborates as an "ecology" of world literature from antiquity to the present. It was the challenge of conducting formal analysis on such a scale that led Moretti to undertake the series of computational experiments he characterized as "distant reading," an approach that seemed to bring literary studies (kicking and screaming) into the age of "big data," its human-scale canons of masterworks subsumed into digitized corpora containing tens or hundreds of thousands of texts.

A "crisis of largeness," then: literary studies driving itself to such extreme scales of analysis that it becomes methodologically unmanageable and incoherent as a regime of value. That is one way to describe the state of affairs in 2016, but it is not the perspective we bring to this special issue. Among other things, we want to take into account the fact that many recent initiatives have sought to conduct literary study on humbler rather than grander scales. Noting the spread of such approaches "across a range of recent literary and cultural study," Mark Seltzer (2011: 727) proposes the rubric of an "incrementalist turn" to capture the shift of

attention and value “toward the minor and the scaled-down.” As surveyed by Heather Love in this issue, the notion of incrementalism links such projects as Sianne Ngai’s (2007: 6) effort in affect studies to revalue “minor and generally unprestigious feelings” with Lauren Berlant’s (2011) affective probing of minor archives, art works, and genres in her scathing account of the historical present; Alex Woloch’s (2003) work on the importance of “minor characters” in the novel; and the several strains of what Stephen Best and Sharon Marcus (2009) label “surface reading”: among them Marcus’s own method of “just reading,” Timothy Bewes’s (2010) “reading with the grain,” and D. A. Miller’s (2010: 126; 2003: 58) “too close reading” (close reading as “humbled . . . minoritized” practice). These developments are not simply reaction formations pushing back against the drive toward larger scales of literary analysis. They stand in a more complex relationship to that tendency, converging with it at numerous points. The rich critical literature that has sprung up around minor characters, for example, has been bolstered by quantitative computational analysis of what Woloch (2003: 13) calls the “character space.” Because it relies on computationally generated quantitative information, this kind of work is too often thought of as committed to the scale of big data. But insofar as it involves close scrutiny of the individual words that make up a particular text, a network analysis of a Shakespeare play (Moretti 2011) or a Dickens novel (Sack 2013), or even Eric Bulson’s (2014) computational book-history project “*Ulysses* by Numbers,” bears as much affinity to traditional fine-toothed study of canonical works as it does to any program of “distant” reading. Here, as in many vectors of contemporary research, we find our discipline wrestling with both the macro and the micro, seeking better, often but not always quantitative ways to coordinate the one with the other. We are experiencing a much more concerted effort than occurred in the late twentieth century actually to recalibrate the entire analytic apparatus of literary study, bringing into view matters of concern both smaller and larger than those whose value we have been accustomed to recognize.

This process may seem to pose a disciplinary crisis in part because new ways of thinking about scale have typically been framed less as ideas than as technologies. “Digital humanities,” in particular, is notoriously difficult to understand as an approach or a school of thought. We can, and probably should, define the phrase by pointing to specific academic

networks (Kirschenbaum 2012). But on its face, the term unifies these diverse teams and projects by grounding them all in digital computers. The phrase *big data* likewise gestures less toward innovative ideas about literary inquiry than toward a set of things: spreadsheets, storage devices, digitized corpora. Terms like these make it easy to understand recent intellectual history as a confrontation between scholars and a collection of technological objects that have intruded on their meditation. “Algorithms,” “data,” and “tools” are often personified, more than half seriously, as historical actors or symbolic antagonists for the humanities (Marche 2012).

In reifying recent developments, technological metaphors also miss their historical specificity. There is nothing terribly new, after all, about data and algorithms as such. Databases and search algorithms were already quietly central to literary scholarship back in the 1990s. The important changes of the last decade are contained not in computers themselves but in a network of unexpected connections to human actors in other disciplines, and to the ideas they are advancing.

A wide array of disciplines is involved in this conversation about opportunities and challenges of scale. The new planetary geographies and temporalities of literary history, for example, are emerging—often in tandem with the most thoughtfully local, place-based forms of “micro-specification” (Cronin 2012)—from environmental studies, in which ecocriticism and the other strains of environmental humanities share intellectual space with the physical and biological sciences, the information sciences, politics, philosophy, economics, and the law. The essays in this issue do not range that widely, but their paths of interdisciplinary exchange are too various to proceed from any single source. What we would emphasize is the part played in this more fluid configuration of disciplines by sociology and other social sciences. Their increasing significance to literary studies is commonly effaced by popularizing accounts that reduce questions of scale to a confrontation between humanists and big data. In gesturing toward computer and information science, such accounts are not entirely wrong. Advances in those disciplines have indeed helped make the whole conversation possible. But few projects in literary studies (and only two essays in this issue) engage computer science directly. We suspect that recent advances in computer science have mattered for literary scholarship mostly in an indirect

fashion, by blurring the traditionally sharp boundary that separated us from the quantitative social sciences.

Literary scholars have conversed with social scientists, of course, for as long as the social sciences have existed. But for much of the twentieth century, literary scholars read social scientists to borrow their conclusions rather than their methods (English 2010: xiii–xiv). Concepts like “habitus” were useful in literary study; the (often quantitative) methods used to develop those concepts seemed less applicable. The subfield of stylistics was a lonely exception: an unusual zone of overlap where literary scholars borrowed quantitative methods from linguists. When computers first made waves in the humanities in the 1960s, they were adopted especially by literary scholars working in this tradition. Writers who want to trace a long, specifically humanistic genealogy for contemporary digital projects often emphasize their continuity with this precedent (Hockey 2004).

The continuity is real, but at this point it represents only one strand in a complex braid of connections, because the intellectual shifts of the last fifteen years have created a host of novel methodological exchanges between the humanities and the social sciences. In fact, it is increasingly difficult to know which discipline is borrowing methods from which. It may be more accurate to say that scholars from a wide range of backgrounds have converged on new research opportunities. An article written collaboratively by a literary scholar and a computer scientist may appear in a special issue of a journal edited by sociologists, along with articles by historians and political scientists (see, e.g., Jockers and Mimno 2013, in Mohr and Bogdanov 2013). Although all these researchers have been drawn together by new methods that involve numbers, the effect has not necessarily been to increase the weight of the things we used to call “quantitative evidence.” On the contrary, new methods have often allowed social scientists to rely more heavily than they are accustomed to on unstructured text (Grimmer and Stewart 2013). So while literary scholars perceive this exchange as tugging them in the direction of social science, social scientists are having the converse experience. In fact, a trio of sociologists has invoked the literary theory of Kenneth Burke to explain the “computational hermeneutics” they see emerging (Mohr, Wagner-Pacifi, and Breiger 2015).



A full history of this conversation would need to describe a series of intellectual and social changes that made it possible. New sources of evidence are one factor. The rise of Google Books in 2004, for instance, created an opportunity for literary study roughly parallel to the opportunities that social media have created for sociologists (Jones 2014: 8; Savage and Burrows 2007). These developments are very much in the public eye and tend to reinforce the perception that recent intellectual history is fundamentally a story about technological change. It is not entirely a false perception: Twitter and Google Books are, indeed, massive new institutions brought into being by the Internet. But the story has other dimensions equally important for scholarship, less widely publicized, and based on conceptual rather than strictly technological advances. The border between the humanities and the social sciences is becoming easier to traverse not just because we have more data now but because quantitative methods themselves have changed.

In the twentieth century statistical models were best suited to describing relations among a few variables. Such models could express a correlation between health and parental income, say, but could do little with a novel (or even a policy brief). Sociologists did often try to use text to produce quantitative evidence, but their methods of “content analysis” required a researcher to define a checklist of topics in advance (“race,” “voting and elections,” “health and medicine,” etc.) and were for that reason usually limited to representing a document’s explicit conceptual content.

The methods now spreading through the social sciences and humanities represent text at a lower linguistic level, using a larger number of variables. The flexibility of this strategy has encouraged social scientists to aspire beyond the merely “semantic” meaning of their sources toward a “poetic” level of interpretation that reflects kinds of meaning implicit in a text’s formal and stylistic dimensions (Mohr, Wagner-Pacifici, and Breiger 2015). How effectively that aspiration can be realized is still open to debate. But in any case literary scholars may be interested to know that these methods have a philosophical foundation that can be engaged and critiqued.

The phrase *big data* tends to suggest that scholars are discussing a purely technical advance, facilitated perhaps by bigger disks or faster computers. From that assumption it is an easy step to the conclusion that

quantitative methods are untheorized, or pit raw data against theory, as an article in *Wired* notoriously claimed (Anderson 2008). The truth is closer to the reverse. The quantitative advances that might matter for humanists had little to do with supercomputers. They emerged instead from forms of epistemological reflection that we would find familiar—if scientists didn't tend to express their epistemological reflections mathematically. Late twentieth-century quantitative models were limited to a few variables, for instance, not because the era's computers couldn't handle more but because models with too many variables succumbed to a form of false precision called "overfitting." A model with more variables than data points can effectively memorize the location of every point, so its explanation seems to fit the data perfectly. But if you present the model with new examples, it may not describe them well at all. It hasn't learned anything generalizable. Researchers in machine learning grappled with this problem by trying to understand what it means to generalize from evidence (Kulkarni and Harman 2011; Vapnik 1999). That inquiry fostered a "new culture" of modeling (Breiman 2001)—a culture that learned to adjust the blurriness of models and test them by asking whether they could generalize from one subset of evidence to another. Statistical methods can grapple with complex textual evidence today not because computers are faster but because human beings have thought carefully about the ways that these models do or don't represent the world.

The value of these methods for literary scholars still needs to be demonstrated, and this will not be an easy task. Although social scientists may see new methods as a step toward the "full hermeneutic complexity and nuance" of their texts (Mohr, Wagner-Pacifici, and Breiger 2015), they are still (by literary standards) incapable of capturing the subtleties of any single work. While our discipline is now more than ever open to experimenting with large-scale analysis, most such experiments will continue to place high value on information gleaned from close reading of individual texts. Even literary historians who contemplate extremely long timelines will tend to default to the logic of the case study, illustrating general trends through rich description of selected particulars. As James Chandler (1998) shows, this approach to history is one of our most enduring orthodoxies. There is little likelihood that particularized case studies and close readings will lose their place of high privilege in the discipline anytime soon.

Quantitative methods stand to contribute most to literary studies when they complement these older practices, filling in scales of description or kinds of interpretive insight that a selection of case studies can miss. Numbers have long been useful for this reason in book history (Altick 1957; St Clair 2004). But that inquiry, of course, focuses on books as objects, seeking historical tendencies in such things as circulation figures or numbers of editions. Is it possible to reveal similar macroscopic trends through quantitative analysis of literary works themselves? This is a philosophical question about what numbers can or cannot achieve but also, more important, a substantive question about existing accounts of literary history. Case studies will need to be complemented by quantitative surveys only if there are actually significant patterns in literary history that scholars have failed to describe. On that question the jury is still out, but one recurring argument in the essays gathered here is that our existing methods do have large blind spots.

New methods, finally, also raise questions about rhetorical effectiveness. Many heuristics that were useful in the research process become encumbrances in the limited space of an article. Truth be told, almost all literary scholars are already applying complex quantitative methods to language when using search engines for discovery. But we suppress that aspect of our research when we describe it. This creates some distortion: the five examples supporting your thesis would be less impressive if you mentioned that they were culled from a population of fifty thousand. But rhetorically, suppression of detail is often the right choice. Our histories of the sonnet form would lose their compressed energy if we had to interrupt ourselves to explain search algorithms. These rhetorical limits pose a challenge for any project exploring literary history on a macroscopic scale, because large scales of description create large numbers of loose ends. But rhetorical compression will pose a particularly severe test for writers who need to unpack unfamiliar methods. Social scientists may be satisfied with accuracy, but literary-critical arguments fail if they fall beneath disciplinary standards of *sprezzatura*. For this reason above all, it is risky to lean heavily on complex methods drawn from other disciplines. It is possible that the empirical methods that matter most for literary scholars will continue to be the simpler forms of counting that have already proved themselves in book history and the sociology of reading (Radway 1984), perhaps assisted by computers but not unduly encumbered by them.

In short, there are new approaches to scale that extend across a spectrum from book history to cultural sociology to practices of statistical modeling shaped by machine learning. There are valid reasons to doubt that the more recondite methods will be widely adopted in literary studies. But our skepticism should at least be based on an informed understanding of new ideas, and aimed roughly in their direction. The temptation to describe a grand cultural struggle between the humanities and the sciences is strong, and humanists' sources of information about science are often out of date. As a result, arguments that purport to critique new methods are often trained in reality at intellectual positions occupied by Karl Popper (1934). Bernhard Rieder and Theo Röhle (2012: 72) speak for many sympathetic but concerned observers, for instance, in worrying that computers may be attractive because they appeal to an old fantasy about "mechanical objectivity"—the belief that the production of knowledge had to be purged entirely from subjective judgement." That would certainly be an odd premise for humanists to adopt. But it is also a premise that could not long survive contact with contemporary practices of computational modeling—which insist, for instance, on Bayesian priors, subjective beliefs that precede observation and are "updated" rather than falsified by new evidence. Unfortunately, disciplines have Bayesian priors about each other that get updated very slowly. Some scientists still believe that the humanities exist to uphold correct grammar, and some humanists still associate any use of numbers with a long-outdated positivism.

Anxieties about interdisciplinary influence can be more nuanced, admittedly, than this summary suggests. Critics who acknowledge that the sciences have developed their own sort of hermeneutic subtlety may nevertheless worry that the subtlety will be lost in translation if quantitative methods are adopted too enthusiastically by other disciplines. (This is roughly the warning that Rieder and Röhle provide, and it is of course valid: scholars should borrow only methods that they genuinely understand.) Other writers suggest that "the cultural authority of digital technology" (Drucker 2012: 85) will make any borrowing of quantitative methods a slippery slope, infecting the humanities with alien premises impossible to contain once allowed in. These concerns are fundamentally predictions, and it is difficult to prove or disprove claims about the future. Perhaps the only way to reply to this dark prediction is

to dramatize an alternate future, where quantitative methods have a place at the table in literary study—facilitating, especially, conversations with social science—without occupying all the places at the table or devaluing the interpretive strengths of other methods.

That vision—not of one discipline in crisis but of several in conversation—is what we try to exemplify in this special issue. All the essays grapple with the relation between scale and value in literary history. While nearly all of them develop points of contact with sociology, they ply different interdisciplinary trade routes and discover different analytic challenges. To provide a context for these present-day rethinkings of scale, Sharon Marcus offers a detailed description of one of the early and most prestigious works of literary history, Erich Auerbach's *Mimesis* (1953). Calling Auerbach the "Alice in Wonderland" of literary studies, Marcus explains his ability to "[vault] from the tiny to the gigantic and back again" as the product of an interlocking set of "critical techniques" derived in part mimetically, from literary works themselves. *Mimesis*, Marcus argues, has retained its high value in literary studies and warrants our special consideration today because it models a strategy, at once intricately technical and intrinsically literary, for managing difficult but necessary relays between the small scales of description and the extended scales of interpretation.

Several essays also address the question of who acquires prestige in the literary world and how this fundamental form of value is adjudicated and distributed. Ted Underwood and Jordan Sellers use machine learning to compare two sets of texts drawn from a digitized corpus of tens of thousands of volumes of nineteenth- and early twentieth-century English-language poetry: a "prestige" set representing volumes that attained a relatively high level of critical regard and a comparison set representing volumes mostly ignored by the cultural elites. The startling success of their algorithmic model in predicting which works were favored by critics across the entire historical span of the corpus leads Underwood and Sellers to conclude that the standards governing literary distinction may be far more stable over time than historians and sociologists of literature usually imagine.

Hoyt Long and Richard Jean So use a computational model not to capture the bases of critical distinction generally but to home in on one particular mode of writing that acquired special value during one period

of time. The object of their inquiry is stream of consciousness, which in the aftermath of European high modernism came to be seen as a marker of the more advanced literatures in the world literary system. Astonishingly, Long and So's algorithm transcends the cultural and linguistic gaps between novels in English and in Japanese. Thus they can perform a comparative historical analysis whose contours confirm Moretti's "wave" model of world literary dispersion. But their analysis also discovers within this wavelike pattern a significant margin of indeterminacy and variance. World literature, they argue, is a product of "turbulent flow," a lively interplay of structure and variance that new computational models are ideally suited to explore.

In comparison with the two essays that precede it, Günter Leypoldt's contribution offers a more guardedly quantitative approach to the social distribution of literary prestige. Taking the works of Walter Scott and Toni Morrison as case studies, he tracks these works' unfolding "social lives": their rising and falling positions in the various spaces where literature is made use of and acquires its cultural traction. Leypoldt's sociological model of the literary field opens up a terrain of "public relevance" in between the zones of critical esteem and commercial success, positing at the same time a new, cultural-sociological scale of literary analysis in between the microhermeneutics of close reading and the macrostatistics of quantitative book history.

Focusing on the recent history of the novel, James F. English confronts different obstacles to the quantitative study of prestige and commerce: copyright law makes it difficult to gather and share large data sets, and the output of the publishing industry now dwarfs even the largest available corpora. His essay highlights a further problem as well: one of the most consequential developments on the field of contemporary fiction—a dramatic reevaluation along the axes of both status and money—pivots around a particular feature of the novel, temporal setting, that has thus far eluded algorithmic detection. English's argument is not that scholars of contemporary literature should disdain quantitative methods but that—for some research problems, at least—we need to build quantitative sociological models on a "'middling' scale," large enough to capture statistically meaningful tendencies of the field as a whole but small enough to be constructed without heavy reliance on computation.

Heather Love steers toward a still finer scale of analysis, revisiting debates in which the most closely observed and minutely detailed modes of social description—whether in realist literature or in interactionist sociology—have been dismissed as “small change,” incapable of addressing larger concerns of politics, power, and “the injustices of the world.” Love reads Claudia Rankine’s *Citizen* as an exemplary work in which an “attention to the microscale” reminiscent of Erving Goffman’s observational sociology proves its value (contra Goffman) as a “political resource.” Through her meticulous accounts of microaggression, Rankine not only draws the smallest acts of everyday racism across the threshold of visibility but locates them as “point[s] of articulation in a larger circuit of violence.” *Citizen* thus models a politically enabling partnership of microsociological practice with literary realism and the documentary lyric.

Finally, Mark McGurl considers a newly dominant institutional agent of literary production: the unprecedentedly large-scale amalgamation of retailer, publisher, library, book club, and literary review that is Amazon.com Inc. Focusing on Amazon’s customer-driven business model and especially on its Kindle Direct Publishing division, McGurl here extends his ongoing project of literary sociology, positing “the Age of Amazon” as “a possible successor-formation to the Program Era,” a new structure for harnessing the impulse to literary creativity on a mass scale. The rearrangements that Amazon is effecting in America’s popular creative culture involve radically new terms of commerce, of course, and new standards of authorial success. But McGurl sees the Age of Amazon as ultimately conserving the presumed relationship between literary value and the longest spans of time, renewing, for authors of the present, the ancient aspiration to “immortality.”

Though these essays share an openness to sociological thinking and a particular concern with matters of scale and quantity, only a few deploy quantitative methods or rely on numbers as such. None of them argues that one can represent literary history only by assembling the largest possible collection of texts in order to cover the past exhaustively. Instead, these essays all rely on some kind of sampling strategy, ranging from a single case study to a selection of several hundred or several thousand volumes. Although operating on very different scales, they are in that sense engaging the problem of scale in parallel ways.

In other words, although these essays rely on different methods and express different opinions, we do not present them primarily as a debate that dramatizes a clash between incompatible (close or distant, deep or surficial) approaches to reading—yet another skirmish in the “reading wars” (Hensley 2012). We would emphasize, instead, a continuity that underlies their methodological diversity. Different scales of literary analysis do require different methods, but if those scales of analysis are inquiring about different aspects of a single social process, there is no reason why the methods should be incompatible. Long-term trends will often make sense only after they are illuminated by local histories; a statistical model, conversely, may reveal a pattern that connects close readings of widely separated passages. We don’t mean to imply that every work of literary history should practice all these methods at once. Constraints of space, not to mention time, make that impossible. But we do mean that it is now possible to leave the reading wars behind, accept the coexistence of different approaches to literary history, and move forward to a stage of this conversation where we ask how different approaches can be productively combined.

---

**James F. English** is John Welsh Centennial Professor of English at the University of Pennsylvania, where he directs the Penn Humanities Forum and the Price Lab for Digital Humanities. He is author of *The Economy of Prestige: Prizes, Awards, and the Circulation of Cultural Value* (2005) and *The Global Future of English Studies* (2012).

**Ted Underwood** is professor of English at the University of Illinois at Urbana-Champaign and author, most recently, of *Why Literary Periods Mattered: Historical Contrast and the Prestige of English Studies* (2013). His essay “Historical Difference as Immortality in the Mid-Nineteenth-Century Novel” appeared in the December 2002 issue of *MLQ*.

## References

- Altick, Richard D. 1957. *The English Common Reader*. Chicago: University of Chicago Press.
- Anderson, Chris. 2008. “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete.” *Wired*, June 23. [www.wired.com/2008/06/pb-theory](http://www.wired.com/2008/06/pb-theory).
- Beecroft, Alexander. 2015. *An Ecology of World Literature: From Antiquity to the Present Day*. London: Verso.
- Berlant, Lauren. 2011. *Cruel Optimism*. Durham, NC: Duke University Press.



- Best, Stephen, and Sharon Marcus. 2009. "Surface Reading: An Introduction." *Representations*, no. 108: 1–21.
- Bewes, Timothy. 2010. "Reading with the Grain: A New World in Literary Studies." *differences* 21, no. 3: 1–33. DOI:10.1215/10407391-2010-007.
- Boyd, Brian, Joseph Carroll, and Johnathan Gottschall. 2010. *Evolution, Literature, and Film*. New York: Columbia University Press.
- Breiman, Leo. 2001. "Statistical Modeling: The Two Cultures." *Statistical Science* 16, no. 3: 199–231. [projecteuclid.org/euclid.ss/1009213726](http://projecteuclid.org/euclid.ss/1009213726).
- Brooks, Cleanth, and Robert Penn Warren. 1976 (1938). *Understanding Poetry*. 4th ed. Boston: Wadsworth.
- Buell, Lawrence. 2007. "Egoblast Affects: The Emergence of U.S. Environmental Imagination on a Planetary Scale." In *Shades of the Planet: American Literature as World Literature*, edited by Lawrence Buell and Wai Chee Dimock, 227–48. Princeton, NJ: Princeton University Press.
- Bulson, Eric. 2014. "Ulysses by Numbers." *Representations*, no. 127: 1–32. DOI: 10.1525/rep.2014.127.1.1.
- Casanova, Pascale. 2005. *The World Republic of Letters*, translated by M. B. DeBevoise. Cambridge, MA: Harvard University Press.
- Chandler, James. 1998. *England in 1819: The Politics of Literary Culture and the Case of Romantic Historicism*. Chicago: University of Chicago Press.
- Cronin, Michael. 2012. *The Expanding World: Towards a Politics of Microspeciation*. Alresford: Zero.
- Damrosch, David. 2011. "Hugo Metzl and the Principle of Polyglotism." In *The Routledge Companion to World Literature*, edited by Theo D'haen, David Damrosch, and Djelal Kadir, 12–20. New York: Routledge.
- Dimock, Wai Chee. 2006. "Scales of Aggregation: Prenational, Subnational, Transnational." *American Literary History* 18, no. 2: 219–28.
- Doyle, Laura. 2015. "Inter-imperiality and Literary Studies in the Longer *Durée*." *PMLA* 130, no. 2: 336–47.
- Drucker, Johanna. 2012. "Humanistic Theory and Digital Scholarship." In *Debates in Digital Humanities*, edited by Matthew K. Gold, 85–95. Minneapolis: University of Minnesota Press.
- English, James F. 2010. "Everywhere and Nowhere: The Sociology of Literature after 'the Sociology of Literature.'" *New Literary History* 41, no. 2: v–xxiii.
- Felski, Rita. 2008. *Uses of Literature*. Oxford: Wiley-Blackwell.
- . 2011. "Context Stinks!" *New Literary History* 42, no. 4: 573–91.
- Greenblatt, Stephen. 1997. "The Touch of the Real." *Representations*, no. 59: 14–29.
- Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis for Political Texts." *Political Analysis* 21, no. 3: 267–97.
- Hayot, Eric. 2013. "Scale, Data, and World Literature." [erichayot.org/wp-content/uploads/2013/11/scaledata.pdf](http://erichayot.org/wp-content/uploads/2013/11/scaledata.pdf) (accessed March 17, 2016).

- Heise, Ursula K. 2014. "Comparative Literature and the Environmental Humanities." 2014–2015 Report on the State of the Discipline of Comparative Literature. American Comparative Literature Association. March 9. [stateofthediscipline.acla.org/entry/comparative-literature-and-environmental-humanities](http://stateofthediscipline.acla.org/entry/comparative-literature-and-environmental-humanities).
- Hensley, Nathan K. 2012. "Figures of Reading." *Criticism* 54, no. 2: 329–42.
- Hockey, Susan. 2004. "The History of Humanities Computing." In *A Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth, 3–19. Oxford: Blackwell.
- Jameson, Fredric. 1981. *The Political Unconscious: Narrative as a Socially Symbolic Act*. Ithaca, NY: Cornell University Press.
- Jockers, Matthew, and David Mimno. 2013. "Significant Themes in Nineteenth-Century Literature." *Poetics* 41, no. 6: 750–69.
- Jones, Steven E. 2014. *The Emergence of the Digital Humanities*. New York: Routledge.
- Kirschenbaum, Matthew. 2012. "Digital Humanities As/Is a Tactical Term." In *Debates in Digital Humanities*, edited by Matthew K. Gold, 415–28. Minneapolis: University of Minnesota Press.
- Kulkarni, Sanjeev, and Gilbert Harman. 2011. *An Elementary Introduction to Statistical Learning Theory*. Hoboken, NJ: Wiley.
- Liu, Alan. 2008. *Local Transcendence: Essays on Postmodern Historicism and the Database*. Chicago: University of Chicago Press.
- Marche, Stephen. 2012. "Literature Is Not Data: Against Digital Humanities." *Los Angeles Review of Books*, October 28. [lareviewofbooks.org/essay/literature-is-not-data-against-digital-humanities](http://lareviewofbooks.org/essay/literature-is-not-data-against-digital-humanities).
- Miller, D. A. 2003. *Jane Austen, or the Secret of Style*. Princeton, NJ: Princeton University Press.
- . 2010. "Hitchcock's Hidden Pictures." *Critical Inquiry* 37, no 1: 106–30.
- Mohr, John, and Petko Bogdanov, eds. 2013. "Topic Models and the Cultural Sciences." Special issue, *Poetics* 41, no. 6.
- Mohr, John, Robin Wagner-Pacifici, and Ronald L. Breiger. 2015. "Toward a Computational Hermeneutics." *Big Data and Society*, July–December, 1–8. [bds.sagepub.com/content/2/2/2053951715613809](http://bds.sagepub.com/content/2/2/2053951715613809).
- Moretti, Franco. 2011. *Network Theory, Plot Analysis*. May 1. Stanford Literary Lab Pamphlet Series. [litlab.stanford.edu/LiteraryLabPamphlet2](http://litlab.stanford.edu/LiteraryLabPamphlet2).
- . 2013. *Distant Reading*. London: Verso.
- Ngai, Sianne. 2007. *Ugly Feelings*. Cambridge, MA: Harvard University Press.
- Popper, Karl. 1934. *Logik der Forschung: Zur Erkenntnistheorie der modernen Naturwissenschaft*. Tübingen: Siebeck.
- Radway, Janice. 1984. *Reading the Romance*. Chapel Hill: University of North Carolina Press.
- Richards, I. A. 1929. *Practical Criticism*. London: Kegan Paul.
- Rieder, Bernhard, and Theo Röhle. 2012. "Digital Methods: Five Challenges." In *Understanding Digital Humanities*, edited by David M. Berry, 67–84. Basingstoke: Palgrave.

- Sack, Graham Alexander. 2013. "Character Networks for Narrative Generation: Structural Balance Theory and the Emergence of Proto-narratives." In *2013 Workshop on Computational Models of Narrative*, edited by Mark A. Finlayson, Bernhard Fisseni, Benedikt Löwe, and Jan Christoph Meister, 183–97. Open-Access Series in Informatics, 32. Wadern: Schloss Dagstuhl–Leibniz-Zentrum für Informatik. DOI:10.4230/OASICS.CMN.2013.183.
- Savage, Mike, and Roger Burrows. 2007. "The Coming Crisis of Empirical Sociology." *Sociology* 41, no. 5: 885–99.
- Seltzer, Mark. 2011. "The Official World." *Critical Inquiry* 37, no. 4: 724–53.
- Shklovsky, Victor. 1990. "Art as Device." In *Theory of Prose*, translated by Benjamin Sher, 1–14. Elmwood Park, IL: Dalkey Archive Press.
- St Clair, William. 2004. *The Reading Nation in the Romantic Period*. Cambridge: Cambridge University Press.
- Underwood, Ted. 2013. *Why Literary Periods Mattered: Historical Contrast and the Prestige of English Studies*. Stanford, CA: Stanford University Press.
- Vapnik, Vladimir N. 1999. "An Overview of Statistical Learning Theory." *IEEE Transactions on Neural Networks* 10, no. 5: 988–99.
- Williams, Raymond. 1989 (1971). "Programmes and Sequences," repr. in *Raymond Williams on Television: Selected Writings*, edited by Alan O'Connor, 133–36. London: Routledge.
- Woloch, Alex. 2003. *The One vs. the Many: Minor Characters and the Space of the Protagonist in the Novel*. Princeton, NJ: Princeton University Press.