



PROJECT MUSE®

1 Gender and Cultural Analytics: Finding or Making Stereotypes?

Published by

Gold, Matthew K. and Lauren F. Klein.

Debates in the Digital Humanities 2019.

University of Minnesota Press, 2019.

Project MUSE. muse.jhu.edu/book/65227.



➔ For additional information about this book
<https://muse.jhu.edu/book/65227>

Gender and Cultural Analytics: Finding or Making Stereotypes?

LAURA MANDELL

Feminist cultural theorists, and perhaps cultural theorists more generally, are discouraged from exploring cultural analytics, or “the computational study of culture,”¹ whenever data miners state their conclusions in terms of “males” and “females.”² In the fields of cultural analytics, computational linguistics, and quantitative sociology, such work typically appears in articles with “gender” in the title—paradoxically because “male” and “female” are biological sex terms rather than gender terms.³ As a keyword for academic sociological and psychological articles, “gender” rose in popularity starkly over the 1980s, from 50 articles between 1966 and 1970 to 1,450 articles between 1980 and 1985 (Haraway, “Gender,” 57). It was of a political moment: as Donna Haraway tells us, gender was “a concept developed to contest the naturalization of sexual difference in multiple arenas of struggle” (53) by feminist writers of the 1980s, culminating in Joan Scott’s seminal 1986 essay, “Gender: A Useful Category of Historical Analysis.” Gender is historicizable because it is culturally constructed. When psychoanalyst Robert Stoller first coined the term “gender identity” in 1967, he contended “that sex was related to biology (hormones, genes, nervous system, morphology), and gender was related to culture (psychology, sociology)” (Haraway, “Gender,” 55).⁴ Therefore, using the biologicistic binary “male” and “female” (M/F) to speak of *gender* is nearly an oxymoron.

Feminist cultural critics might be tempted to simply walk away from cultural analytics altogether rather than engage with what I call “M/F”—that is, the conflation of gender with sex and the naturalization of a culturally constructed binary opposition. But walking away is a bad idea. During the 1980s, feminist sociologists critiqued patriarchal notions of objectivity primarily by turning away from quantitative analysis toward the qualitative (Gorelick, 465; Hartsock, 232). For Joey Sprague and Mary Zimmerman, the failure of feminism to have as transformative an impact on sociology as it has in had “in ‘anthropology,’ ‘history,’ and ‘literary criticism’” resulted from feminist sociologists’ preference for qualitative analysis (84). Quantitative analyses in sociology proceeded without feminist input at great cost, leaving

the field susceptible to “bad science” (Haraway, “Gender,” 55) and “bad description” (Marcus, Love, and Best, 6).

In addition to ensuring that feminist issues are addressed in the field, investigating M/F computational analyses also gives us the opportunity to revisit a blindness pervasive in 1980s gender analysis. For one thing, we now commonly accept that genders are not reduced to only two, nor are they wholly distinct from sex: cisgender, transgender, and genderqueer, for example, designate not masculine or feminine gender identities, but instead the relationship a person has to his or her assigned gender (LaBrada). Furthermore, a person’s sex/gender, as so-called biological sex is now sometimes described, is no longer taken as a given: it is “assigned” to a person by medical and familial institutions. That a person is given a sex/gender after being born is made salient in cases of chromosomal and genital variation, but gender assignment nonetheless happens in every instance. The video montage for the opening credits of Jill Soloway’s comedy-drama web television series, *Transparent*, does an excellent job of displaying women and men more or less easily assuming their assigned gender roles or, alternately, stepping outside them—so much so that, when a clip of a young girl joyously twirling her skirt appears in the montage, one realizes that she is celebrating her own capacity to fully embrace and express her incipient womanhood: hers is a moment of cisgender triumph—that is, a feeling of triumph for an achievement.

This understanding of sex, gender, and their relation owes much to the work of feminists in the 1990s, who extended into thinking about sex categories the work done by feminists in the 1980s to question the role of nature in determining gender categories. Judith Butler, for example, disturbed the boundaries of the sex/nature versus gender/culture opposition in *Gender Trouble* (1990): “Gender ought not to be conceived merely as the cultural inscription of meaning on a pregiven sex (a juridical construction); gender must also designate the very apparatus of production whereby the sexes themselves are established” (7). The “apparatus of production” is for Butler not only a legal and medical intervention. The biological binary M/F opposition is a textual and a numerical one: it is written, printed, spoken—discursive.

Some currently publishing data miners, sociolinguists, and sociologists uncritically conflate sex with gender, regressing to the pre-1980s intellectual milieu in which the M/F categories were treated as designating natural, transparent, biological categories. But some do not. For example, studies by Mark Olsen (“Qualitative Linguistics and *Histoire des Mentalités*: Gender Representation in the *Trésor de la Langue Française*,” 1992) as well as by Ted Underwood and David Bamman (“The Instability of Gender,” 2016), theorize gender as a category in the making, rather than as a preexisting fact. Using data-mining techniques on the ARTFL corpus (texts gathered for the project, American and French Research on the Treasury of the French Language⁵), Olsen has shown that stereotypical associations with the category of woman change over time (“Qualitative Linguistics,” 13). This work by Olsen, as well as that by Underwood and Bamman, demonstrates that computational

methods offer the opportunity to denaturalize gender categories. But to capitalize on this opportunity, we must remain attentive to where and how computational data analytics crosses the line into naïve empiricism—often unwittingly—by presenting conclusions about “male” and “female” modes of thinking and writing as if the M/F terms were simple pointers to an unproblematic reality, transparently referential and not discursively constituted.

My argument in this chapter is that, in some quantitative cultural analyses, the category of gender has been biologized, binarized, and essentialized in a trend that I call “stereotyping,” with an eye on the historical and material practice of stereotyping. Stereotyping as a material practice was invented before the printing process became industrialized (1725 in England, 1735 in France). It froze pages made from movable type into unalterable plates of textual data for reprinting and was adopted early on by the French printing house known as Didot (Hodgson 15–18). The term for “stereotype” in French is *cliché*.⁶ Here, I show how data-mining techniques are sometimes used to produce clichés. But my goal is less to dismiss quantification than it is to argue that, insofar as the results of data mining are movable and manipulable, cultural analytics can serve us well in historicizing gender categories. Examining the gender problem in quantitative analysis reveals both opportunities afforded by digital media and how we might miss them.

Text mining devolves from process into fixed results via stereotyping, whether literally imprinted on paper or figuratively frozen into cliché. In contrast, acknowledging that numbers have to be *read*—that is, resisting the view that statistical results are meaningful in any self-evident way—takes place by close reading of computational results and often manipulating parameters, adjusting algorithms, and rerunning the data. Dynamically reading, generating, and analyzing numbers are essential to making interpretive quantification possible and thus to exploring how the M/F binary is constituted in any given historical moment. This chapter examines two instances of gender analysis, performed via stylometry, word-frequency analysis, and topic modeling, that veer toward naïve empiricism because they take the distinction between M/F to be transparent. Imbuing cultural analytics with feminist theory, I argue here, shows us that we can animate numerical processes rather than fixing their results as stereotype. New technologies and carefully created reading processes can transform the statistical analysis of cultural artifacts into dynamic readerly engagement.

Bias and Statistics

Matthew Jockers undertakes what he calls “macroanalysis,” he says explicitly, in order to strip literary criticism of its interpretive bias. In his book of the same name, he defines macroanalysis as “computational approaches to the study of literature” (4).⁷ Despite later conciliatory gestures, Jockers unequivocally discredits more interpretive literary approaches, including close reading, in the book’s first pages: “Close

reading is not only impractical as a means of evidence gathering in the digital library, but big data render it totally inappropriate as a method of studying literary history” (7). For Jockers, close reading is “totally inappropriate” because it unleashes “irrepressible biases.” Here, he takes a peculiar theoretical step that allows him to implicitly claim that biases can be dismantled by text mining: “Observation is flawed in the same way that generalization from the specific is flawed.” According to this line of reasoning, interpretive generalization fails only as a result of working with too small a sample size (6–7). For example, Ian Watt’s interpretation of the rise of the novel, Jockers says, turns out to be “*correct* on a number of points” despite being performed on too small a number of texts (7, emphasis added). His own assessment, because it is based on more texts, has therefore incontrovertibly established Watt’s thesis, despite many critical studies that contest it.⁸

Before discussing Jockers’s conclusions concerning gender and literary style, I want to look closely at his methods. One is stylometry, which the O.E.D. defines as “the statistical analysis of variations between one writer or genre and another.” In *Macroanalysis*, however, stylometry is used to analyze variations between “male” and “female” writers of novels. Stylometrists typically delineate styles using very small words, not those that are related to the topic: the difference between the styles of James Madison and Alexander Hamilton as visible in the *Federalist Papers*, for instance, hinges on how often one uses “a” and the other “and” (Jockers, 92). Julia Flanders describes the work of John Burrows, another proponent of stylometric analysis, as discovering “distinctions of idiolect” usually invisible to readers; for instance, the frequency with which “‘inert words like ‘of’ and ‘the’” are used (56). At first, Jockers sticks somewhat to the standard method: he describes the two features that most distinguish “the gender signal” as the higher usage of feminine pronouns and of the colon and semicolon. Jockers acknowledges that feminine pronouns are used more in novels written about women, so this stylistic feature is more about the content of novels written by women than about their style. But more problematically, as textual editors well know, punctuation (“accidentals,” in the parlance of textual editing) is often manipulated by editors of print editions. What, we might well ask, were the genders of the editors?

Jockers further deviates from Burrows’s method in producing lists of “features” used by female authors, by which he means specific vocabularies or keywords: “It is also seen that women authors show higher usage of the words *heart* and *love*—not surprising in the sense that this fact is consistent with our stereotypes of ‘feminine’ prose” (*Macroanalysis*, 93). Here, Jockers veers away from stylometry into simple counting. But there is a bigger problem. Not all data miners would attribute to a dataset spanning 200 years “our stereotypes.” This phrase controverts the very point Joan Scott was making by introducing the term “gender” into cultural history: that gender associations change over time.⁹

A study by Jan Rybicki seeks to build on Jockers’s work. Rybicki uses stylometrics to determine the gender of twelve anonymous authors in a collection of fiction

held by the Library at Chawton House, which used to belong to Jane Austen's brother Edward and now proclaims itself to be "home to early women's writing."¹⁰ Rybicki supplements this small corpus of forty-six novels with novels written by canonical authors, including men. He claims that he is able to sort these novels by the gender of their authors by using the same technique that sorted the canonical authors by gender in order to determine the genders of the anonymous authors in the Chawton House dataset. He too relies on the frequency of substantive vocabulary rather than of inert words. His conclusion resembles that of Jockers: "The Chawton House authors exhibit a decided preference for words that would be most readily and stereotypically associated with sentimentalist feminine fiction, and the fact that the stereotype is visible in vocabulary statistics makes it, perhaps, less of a stereotype" (754). Among the gendered characteristics that Rybicki has imagined as "less of a stereotype" is the finding that "men like virtues," by which he apparently means to suggest that men really do always care more about virtue than do women. His vocabulary statistics, however, show women using "virtue" and "virtuous" frequently and both men and women using "honest" (9). More obviously than in Jockers's analysis, the stereotype is in the eye of the beholder.¹¹

Leaving aside Jockers's and Rybicki's reliance on stereotypes to explain their results, it is worth exploring whether the results themselves can ever be taken as objective. Results achieved by Mark Olsen contradict those of Jockers and Rybicki: he finds no such thing as a "feminine style" in his data. Although contradiction itself cannot overturn claims for objectivity, Olsen's method reveals the extent to which Rybicki's and Jockers's results are based on a preexisting assumption. Olsen deliberately does not look at word frequency (*Écriture féminine*). For Olsen, what distinguishes feminine writing is not vocabulary because it is tied to topic. For him, a gendered style consists in using the same words, but in a different way. He therefore examines the contexts of specific words, concluding that the *meanings* of words do not diverge based on gender. For Olsen, then, women and men make deliberate decisions about how to use specific words. Therefore, Olsen was doing something other than counting words according to an author's gender because he believes topic to be the overriding factor, not gender, in determining which terms will be used. Olsen's articulated premise for his study reveals Jockers's and Rybicki's underlying assumption: choices about diction are not the product of writerly decisions related to topic, but instead are based on being either M or F. As will become more and more clear, Jockers's and Rybicki's results confirm not an objective reality, but the premise of their own studies.

The criterion of scientific objectivity is reproducibility.¹² Before finding his own "female" words, Rybicki attempted to sort a subset of his corpus by using Jockers's list of "features best distinguishing male and female authors" (94, qtd. Rybicki, 5) and a second list created by James Pennebaker in a book called *The Secret Life of Pronouns*. Neither list worked; he had to generate a new list of "gender signals" from his own corpus. Rybicki attributes the failure of Pennebaker's wordlist to effectively

sort gender in Rybicki's own corpus to the fact that the two corpora cover different time spans. But he also observes that Jockers's corpus contains "some of the authors in my [canonical] corpus" (5). Rybicki mitigates the failure of Jockers's list by noting certain overlaps with his own (5) and concludes that his own list "makes sense in terms of Pennebaker's and Jockers's findings" (7). But the fact remains that Rybicki could not use either of the two lists to accurately sort his corpus. His attempt to replicate Pennebaker's and Jockers's findings failed.

Clearly, the ability to distinguish gender via particular words depends upon the construction of a particular corpus. Here it is important to notice a moment of clarity in Rybicki's richly contradictory article. It is not simply that there are biases intrinsic to data selection (see also Koolen and van Cranenburgh, 14–15), but that irreproducibility suggests the instability of discursively constructed gender categories: "Literary scholarship has never voiced it as clearly as quantitative research has that the notions of 'writing like a man' and 'writing like a woman' are in such a constant state of flux that they risk becoming, in fact, quite problematic" (14). While one might argue with Rybicki's assessment of "literary scholarship" — Toril Moi's *Sexual/Textual Politics* of 1985 summarized all the work in the field that problematized "writing like a woman" — his claim needs to be trumpeted loudly among data miners: it is "literary evolution" and not bodies that assign keywords to M/F.

Rybicki concludes by noting the "inability, in this study, to produce a stable 'canon' of male and female keywords that would survive a change of corpus or shifts in literary evolution, or both" (14). He proudly attributes this insight to "quantitative research," which is better, he says, than ordinary "literary scholarship" in figuring out that "writing like a woman" changes through time. Why, then, is this claim so inconspicuous as to not even inform the title of his article ("*Vive la différence!*") or the main argument—that he *can* determine the gender of anonymous authors in the Chawton House collection?

Gender difference tends to be widely broadcasted, and anything that obfuscates it downplayed. For instance, *Macroanalysis* made the news because it "shows that, on a whole, female authorship is detectable by objective measures rather than just human intuition" (the Smithsonian's *Smart News*, Nuwer). In her meta-analysis of psychological gender studies, Janet Sibley Hyde argues that the popular press picks up on and touts gender differences, even in studies that conclude there are very few marked differences. Studies of psychological gender traits have, since 1879, been reporting *more similarities than differences*. Most striking, in *The Psychology of Sex Differences*, Eleanor Maccoby and Carol Jacklin argued that most recorded gender differences did not hold up to scrutiny, and yet their work is reported in textbooks and reviews as discovering difference, not similarity. In her analysis of this fact, Hyde observes, "The mass media and the general public are captivated by findings of gender differences" (581).

In her own work, Hyde has advanced what she calls "the gender similarities hypothesis, which holds that males and females [*sic*] are *similar* on most, but not all,

psychological variables” (581, emphasis added). Compiling and comparing many studies of gender difference, Hyde concludes that her gender similarity hypothesis proves true except in the case of some physical differences (throwing distance) and some aspects of sexuality (590). Stereotypical differences such as higher aggression levels and better mathematical capacities turn out to be attributable to men only in supervised tests; that is, when women know they are being watched or tracked for evidence of their gender (589; s.a. Cole, 178). Hyde’s gender similarities hypothesis is consistent with linguist Kate Harrington’s deeper reading of the statistics concerning gender differences in language. Harrington has looked more closely at the numbers generated by gendered discourse analysis, examining rather large differences in standard deviations between those classified in each gender. However, if one leaves out the extreme cases—those who rank most highly in gendered characteristics—most women resemble most men: “the often cited difference [is] dependent on the few and not representative of the many” (96). Major differences in standard deviation between the (allegedly) two genders can indicate that a dataset warrants further investigation, but the standard deviation will not be markedly different if there are exaggerators in both groups: the “*range* of individual variation . . . needs to be interpreted” (96, Harrington’s emphasis). Moreover, Harrington finds that those people who grossly exaggerate their own “stereotypical” gendered discourse are members of specifically identifiable groups. It is the intersection of identities that matter, not gender alone: any “speaker’s use of [allegedly gendered discourse forms] has more to do with individual social context than with gender as a social category” (101). The “individual social contexts” identified by Harrington as displaying the largest gendered differences were teenage girls (97) and men who seemed especially threatened by the study of gender differences in which they were participating (100). “The only overt comment on ‘doing gender’ in my data,” Harrington says, arises in an explicit discussion of masculinity using the now infamous genre of discourse, “Real men don’t do X” (99–100).

In my view, the exaggerators indicate, again, that gender identity is not given, but is triumphantly attained through the course of developmental acculturation (in the case of the teens) and sometimes is an insecure achievement (in the case of the threatened men). People who adopt and exaggerate stereotypically gendered traits (as they understand them) are worried that others might see their gender as ambiguous. That gender panic may intervene in sociological studies offers insight into why differences are seized on and trumpeted by the popular press. Gender is bifurcated into two categories, and people fit themselves into one or the other with difficulty and trepidation. That process requires continuously reassuring ourselves that identifying (as) M or F is no trouble at all. Reportage of firm gender demarcations is an apotropaic gesture designed to ward off anything that undermines a symbolic difference essential to the Western world’s self-understanding.¹³

Given this discussion of gender studies, what can we conclude about the extent to which statistical reading techniques can mitigate irrational symbolic

investments? Statistics, Isabelle Stengers notes, are deeply conservative in how they build on previous results (n. 10, 263). For instance, Bayesian probability, the foundation of latent Dirichlet allocation (a form of topic modeling, another statistical technique, discussed later in this chapter), determines events to be more probable if similar events have already happened. More generally, probability contributes to the biases of computer-assisted racial profiling (Hardt; O'Neil) and even Google search returns (Noble).

However, the discipline of statistics has developed methods for ameliorating bias. Ben Merriman notices Franco Moretti's and Jockers's tendency to avoid using known statistical methods for reading their results. In particular, they rely a great deal, Merriman says, on statistical significance for their arguments, when in fact statistical significance proliferates everywhere one looks for it. If a dataset of novels were categorized according to novelists with long noses and novelists with short noses, we would find statistically significant differences; they simply would not be meaningful differences. Although many researchers turn to qualitative analysis to determine the substantive value of a difference (Fitz-Gibbon, 48), there are also quantitative methods for estimating whether a difference is meaningful or its "effect size" (Coe). The effect size, or "how much difference a difference makes (the magnitude of the effect)" (Fitz-Gibbon, 45), is typically used for "meta-analyses" such as Hyde's, but it need not be used only for comparing multiple statistical results from different studies. To be fair to Moretti and Jockers, determining effect size is neither typical nor easy. As Robert Coe (1) tells us, "despite the fact that measures of effect size have been available for at least 60 years (Huberty, 2002), and the American Psychological Association has been officially encouraging authors to report effect sizes since 1994—but with limited success (Wilkinson et al., 1999)."¹⁴ What the need for measuring effect size indicates, however, is that, as Merriman puts it, "patterns may be statistically significant but substantively trivial: the magnitude of the differences (the 'effect size') may be extremely small, or it could disappear entirely if the researcher tweaks the model" (16). The irreproducibility of gendered keywords confirms Merriman's hypothesis.

But even after determining effect size, identifying statistical trends does not remove bias. On the contrary, it reveals the biases of the paradigm according to which a thing or event is counted as data (Forrester, 36, qtg. Hesse, 286). Statisticians well know that biases are reproduced through classification systems (Cartwright and Bradburn, 5). Comparing the work done by Burrows, Flanders, and Olsen to Jockers and Rybicki (discussed on pp. 6–8 above) reveals a bias toward biological determinism.¹⁵

"Gender" was proposed as an analytic category in the 1980s specifically to avoid the deterministic bias. According to Donna Haraway, it "was developed as a category to explore what counts as a 'woman'" through time ("Gender," 67). But setting aside "as a 'woman'" from Haraway's statement for a moment, *What* counts?¹⁶ The question becomes axiological: What really has value for us, and why? I sketched out earlier some thoughts about the symbolic weight and psychological drama that

obfuscate for us the conceptual inadequacy of the M/F categories. There is also a material side to the redacted question: *What* has been charged with the procedure of counting? It would be very easy to answer, in the case of cultural analytics, that what are responsible for counting are data-mining techniques and data miners. But that answer rings true only if the data itself is not at all actively participating in counting, unlike Harrington's interviewees, who very actively generate a feedback loop. Are there feedback loops built into data? Has it been collected, sorted, and even produced according to the categories M/F? It is to answer those questions, in the case of text-mining novels, that I now turn.

Data as Capta

In a keynote collaborative presentation with Julia Flanders at the Boston Area Days of Digital Humanities Conference in 2013, Jockers explained the difference between the visualizations that he originally generated for *Macroanalysis* and those finally published in the book: in the course of revision, Jockers discovered that there were duplicate novels in his dataset and that three-volume novels had not been stitched together to count as one novel instead of three. Changing his dataset also altered a dramatic visualization of gender difference included in an earlier publication ("Computing and Visualizing," Figure 3), transforming it into the less distinguishing visualization that appears both in *Macroanalysis* (Figure 9.3.4) and in *Smithsonian Smart News* (Nuwer).¹⁷ In other words, the stylistic differentiation between M/F genders no longer looked so significant once Jockers tweaked his dataset.

Writing on the subject of the visual effect of data visualizations more generally (*Graphesis*, 129), Johanna Drucker argues that the "representational force" of the image "conceals what the statistician knows very well—that no 'data' pre-exist their parameterization. *Data are capta*, taken not given, constructed as an interpretation of the phenomenal world, not inherent in it" (128, emphasis in the original). Data-sets are never random samples: they are gathered by humans according to parameters, however complex, hidden, unconscious, and subject to chance they may be.

Drucker's formulation of data as "capta" cuts two ways. First, why collect data according to the categories M/F? Capturing data as M/F will create statistically significant differences between M/F, but are they significant? And more importantly, what exactly do they signify? Second, thinking about data as capta encourages us to examine the selection process in aggregating data. The shape of Jockers's dataset was decisive in determining his results, as is clear from Rybicki's proudly avowed inability "to produce a stable 'canon' of male and female keywords that would survive a change of corpus, or shifts in literary evolution, or both" (14).

It is a virtue of Rybicki's stylometric analysis of the Chawton House Collection of novels that he openly describes the process of tweaking his dataset—changing, over the course of his experiment, which novels by men he would include or exclude—to get his algorithm to successfully differentiate authors marked as M/F.

Rybicki tells us that he removed from his corpus the “troublemakers” who disturbed his results too deeply because it is “notoriously too difficult to gender classify” (3). That is, gender classification does not work without help.

Rybicki’s adjustments are not surreptitious: he is trying to get his algorithm to work through acts “of honest and well-founded cherry-picking” (6). In a 2013 article, Alan Liu quotes a question broached in 2011 by Ryan Heuser and Long Le-Khac: “How do we get from numbers to *meaning*?”¹⁸ Liu argues that this “meaning problem” (“What is the Meaning,” 411) is not simply a matter of interpretation *after* data has been generated, describing abstractly what we can see in the practical way that Rybicki works:

Any quest for stable method in understanding how knowledge is generated by human beings using machines founders on the initial fallacy that there are immaculately separate human and machinic orders, each with an ontological, epistemological, and pragmatic purity Digital humanities method . . . consists in repeatedly coadjusting human concepts and machine technologies until . . . the two stabilize each other in temporary postures of truth that neither by itself could sustain. (“What Is the Meaning,” 416)

Adjusting the dataset to get the best results is simply what data miners do to achieve “temporary postures of truth.” This is good practice, not a crime, and it is *only* problematic insofar as the data miners present their findings as “objective truth,” obscuring to us (and sometimes to themselves) their own role in massaging data and adjusting the parameters of their algorithms to make them work. It would be very easy to break Rybicki’s algorithm, to prevent “highly adventurous content” from being associated with “M” and sentimentality with “F,” simply by adding novels to the dataset.¹⁹ Data amassed, which is to say data, is always capta: counted or captured by someone.

Gender/Genre: What Counts?

To arrive at a working method for assessing the genders of the anonymous authors in his dataset, Rybicki must omit certain other authors from his collection of writers with known genders: this shows us how genre, as much as gender, muddies the picture. Genre affects the algorithm’s ability to distinguish M/F in three ways. First, “feminine” sentimental literature and “masculine” criticism incorporate stylistic features that are discernible and imitable. Rybicki had to eliminate William Godwin, who wrote in the genre of Jacobin novels that are typically sentimental in style. His results also show Samuel Richardson’s epistolary novels, written largely in the voices of his eponymous heroines, veering toward “F.” I asked Jan Rybicki to use his algorithms to perform a stylometric analysis of Mary Wollstonecraft’s two anti-Jacobin, sentimentalist novels along with two essays, her two *Vindications*, in

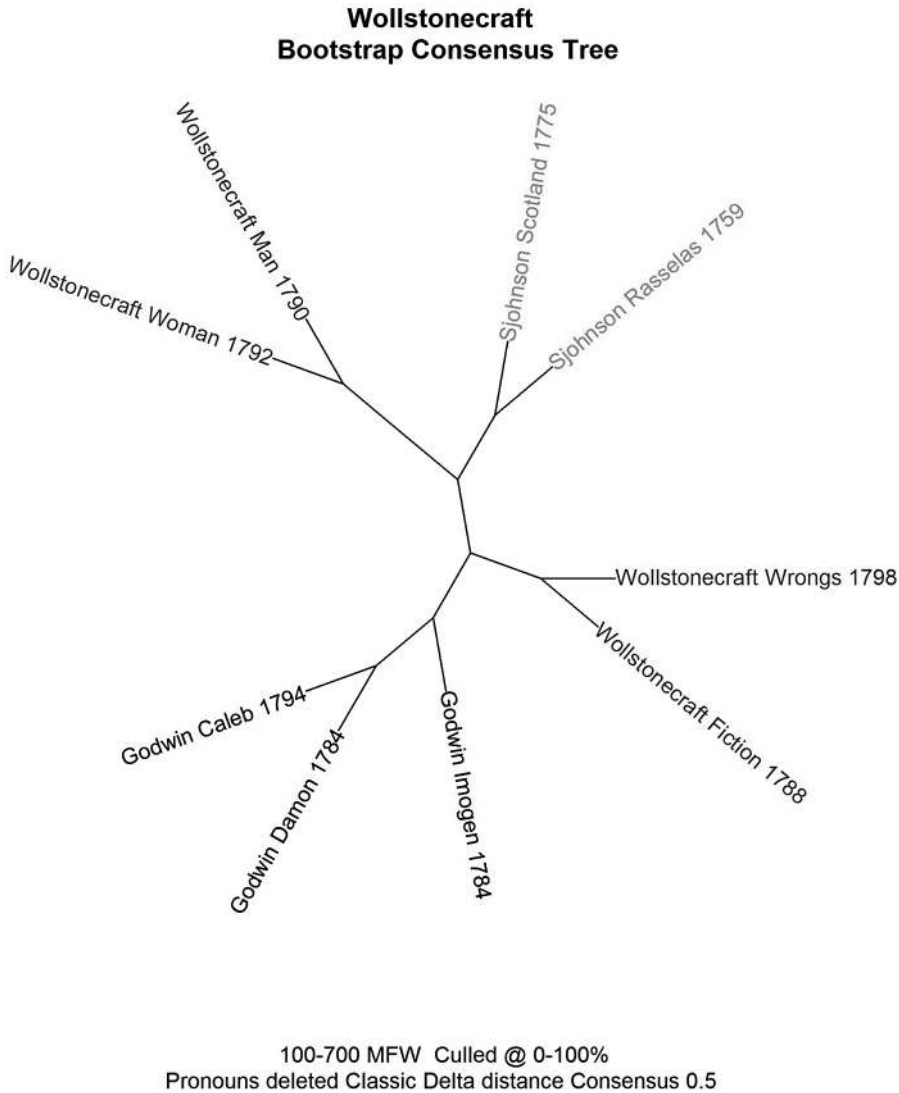


Figure 1.1. This bootstrap consensus tree groups together texts with similar styles using stylometry; specifically, the Delta measure of stylistic difference pioneered by John Burrows that pays attention to the use of most frequent, little words (“a,” “of,” “the”), as well as keywords. Wollstonecraft’s essays are farthest away from everything else, including her own novelistic writing, as indicated by the double-length line. *Rasselas* might be technically classified as a novel or at least a fable, but it is distinctive because of Johnson’s style. *Rasselas* excepted, the biggest divide shown here is between novels and essays, not M/F. Thanks to Jan Rybicki for running this test.

relation to texts by two men. Both of Wollstonecraft’s essays are, to my sensibility, very Johnsonian in style, characterized by Latinate diction and complicated syntax. I asked him to test at least in one case the effect of genre on gender differentiation, and he produced the result shown in Figure 1.1.

Wollstonecraft's sentimental anti-Jacobin novels most resemble Godwin's sentimental anti-Jacobin novels, as indicated by proximity in this bootstrap consensus tree, whereas her essays most resemble Johnson's writings. In fact, Wollstonecraft chose to write both her *Vindications* in the Latinate, Johnsonian style that dominated the essay genre because it had, at her moment, considerable cultural authority. The literary critic Nancy Miller points out that two other eighteenth-century novelists, Frances Burney and Françoise d'Issembourg d'Happoncourt Graffigny, assume a similar style when, in prefaces to their novels, they take on the style of a critic (51–52). Insofar as one pins a genre to a gender, then a woman writing in a Latinate style (essays, criticism) or a man writing a Jacobin sentimental novel is writing in drag.

The case of genre helps us understand more about gender analysis; that is, about figuring out “what counts as ‘woman’” (Haraway, “Gender,” 67). It shows us that features of both gender and genre, while highly discernible, are also highly imitable. Writing in the sentimental genre can be accomplished through study. For instance, Rybicki notices one parody of “female writing” by William Beckford (7). Despite his awareness of the possibility of parody, Rybicki argues that one of the anonymous Chawton House novels is “correctly” identified as written by a man because its preface deprecates “‘womanish and weak minds’” (13, emphasis in the original). It could be a parody, but it also could be written by a woman pretending to write as a man. In fact, Wollstonecraft originally published her *Vindication of the Rights of Men* anonymously. In it, she disparages Edmund Burke, whose *Reflections on the Revolution in France* she is attacking, as “effeminate.” That kind of slur certainly helped disguise her gender and with it her identity. Deprecating femininity is a gender marker that women find easy to imitate in order to establish their authority. (We can see from Wollstonecraft's second *Vindication* that she did honestly devalue “effeminacy,” which she analyzed as learned rather than innate, pointing out that the soldiers of her time and class behaved with as much effeminacy as women.) Anyone can adopt gendered modes of behavior, just as anyone can write in genres stereotypically labeled M/F.

In writing, if we define “gender” as a category-in-the-making (Haraway, *Modest_Witness*, 35), as a set of conventions for *self*-representation that are negotiated and manipulated, more or less consciously, then gender approaches nearly to the definition of genre. (If we take out “self” from the preceding sentence's proposed definition of gender, the result, “a set of conventions for representation,” is a definition of genre.) Learning to write in a specific genre is almost as difficult as learning to encode one's self-expression according to a gender—and in some instances, the tasks largely overlap.²⁰ In person, one sends gender signals with body movements and gestures, speech, and writing. Whereas in the case of gender, we think of imitating its forms as “passing” or “drag,” in the case of genre, unless it is a case of parody, we think of it as, well, writing. In composing her two *Vindications*, Wollstonecraft is not imitating an essay writer; she is writing essays. When Rybicki says in his essay that Austen “writes more like a man” (9), maybe that too is only evidence of Austen writing.

Am I arguing, then, that M/F is a bogus binary, produced by data collectors and data miners, without corresponding to any distinction actually found in reality? There are two answers to this question. First, yes: I believe the gender binary is not really about gender. But if it is wrong, how is Jockers's stylometric analysis able to determine a novel's gender correctly in 80 percent of the cases, while identifying genre correctly in only 67 percent? That fact justifies Jockers's statement that "gender . . . is a strong signal" (*Macroanalysis*, 93). However, if we look at the numbers more closely, what Jockers is calling "the genre signal" is weaker than what he is calling "the gender signal" because his algorithm is trying to sort texts into eleven genre categories, whereas there are only two gender categories: M/F. His algorithm is actually eight times better than chance at determining genre, but only 1.6 times better when determining gender.

I would rather roll dice that are 1.6 times more likely to give me a winning number.²¹ But I would maintain that the algorithm sorting texts into M/F is not sorting *only* gender at one moment, and *only* genre at another, when it is required to sort in two categories as opposed to eleven. The gender/genre signals, as Rybicki's work subsequently shows, are inextricably intertwined. In other words, when stylometric analysis guesses 1.6 times better, it is guessing about what I would call "textual gender," the stylistic and textual features associated with gendered genres. It is by now a critical commonplace that gender construction is discursive: it takes place in writing, in everything from medical documents and legal cases to novels. And many, many books and articles examine the relationship between gender and genre. Here, however, I am proposing that, when gender analysis operationalizes data-mining algorithms, separating gender from other markings (genre, era of composition) is not possible: historical time and genre are not incidental to, but constitutive of, gender.

We could break the algorithm's capacity to produce "a strong gender signal" by simply increasing the number of gender categories to be sorted. Experts in the field could create metadata to generate a completely new taxonomy to replace the tired M/F binary: "men writing as men," "women writing as women," "women writing as men," "men writing as women," "unspecified (anonymous) writing as men," "unspecified writing as women," "men writing as men (byline) in the voice of a woman (woman narrator)," "men writing as unspecified (anonymous byline) in the voice of a woman," "women writing as men (byline) in a voice of unspecified," etc.—whatever categories are presented by title pages, prefaces, narrators' discourses, and research into authorship attribution found.²²

Breaking down the strength of a signal is not the same as disrupting computational capacity, which, as Ted Underwood argues, no longer depends on simple classification schemes:

Maybe in 1967 it was true that computers could only handle exclusive binary categories. I don't know: I was divided into a pair of gametes at the time myself.²³ But nowadays data models can be as complex as you like. Go ahead,

add another column to the database. No technical limit will stop you. Make categories perspectival, by linking each tag to a specific observer. Allow contradictory tags. If you're still worried that things are too simple, make each variable a probability distribution. The computer won't break a sweat. ("The Real Problem with Distant Reading")

Computation enables complexity.

Additionally, the computer screen affords interactivity. Text-mining parameters can be manipulated by viewers, as they are in Voyant windows.²⁴ The authors of Voyant, Geoffrey Rockwell and Stéfán Sinclair, have also created *Methodi.ca*,²⁵ through which algorithms can be altered. The new "Data" section of *CA: Journal of Cultural Analytics* allows not only for sharing data but also for articulating the principles of data collection.²⁶ Visualization tools are being built that will allow running and rerunning data, each time changing the underlying parameters used for text mining in order to test their effects, and that also make it easy to drill down to the text itself for close reading when questions arise.²⁷ Digitization forces on us neither rigid quantification systems nor distant reading. Far from it: digital media can foster the mixing of qualitative and quantitative methods and represent the results of analysis from multiple perspectives.

The second answer to the question posed earlier is no: M/F is not a bogus binary. Eighty percent is a very good rate of correctness in data mining. But exactly *what* has been correctly measured? Feminist literary historians have known for a long time that the early novel was indeed a place where gender norms were formed (Miller, 47; Gelfand and Switten, 443). Furthermore, this construction of gender took place at the very same moment, and using the very same literary devices, as the moment and methods through which novelistic genres were formed. As seen in Harrington's analysis of speech patterns described earlier, people partly create and partly confirm stereotypes. Harrington's exaggerators seized on existing classifiers and amplified them. Thus, the M/F classification process provides a feedback loop: people wishing to exaggerate their gender will repeatedly send signals that are then further confirmed after the fact.

As novelists attempted to create gender norms, those norms became more imitable, more easily parodied,²⁸ and therefore more easily adopted throughout the culture. Novelists were counting and reproducing as many gender and genre features as they could. Underwood and Bamman's study of gendered norms in novels written between 1780 and 1989 finds a general trend: over time "gender [becomes] harder and harder to predict." After a certain historical moment, perhaps after the Victorian era (Jockers and Kirilloff), novels turned to tasks other than stereotyping the M/F binary. We could use stylometry to find this out. But stylometry will only become a valuable tool of literary criticism if we admit that categories such as gender are being constructed both by the measurer and the measured. Once this process is accepted, we might then be able to use stylometry to experiment with

new taxonomies of gender, such as the one adumbrated earlier. If trained on taxonomies generated by experts in feminist literary history, who have familiarity with the time period, the guesses that an algorithm might make about gender, although a weaker signal because sorting into a greater number of categories, would optimize human-computer interactions. Algorithmic sorting would sometimes reveal our blindspots, surprising us with insight, and sometimes reveal its own stupidity, marking those moments when human thinking surpasses computational logic. It would never simply be “right” or “wrong.”

In *Modest_Witness*, Haraway protests against what she calls “virile witness,” testimony as an act of opposition. The modest witness, by contrast, though no less biased as a human being, attempts to identify the terms of self-representation without forcing them into the form of aggressive opposition.²⁹ Rewriting gender as sex, wrongly understood as a simple binary opposition, transforms it into a weak signal that might be statistically “correct.” Yet, with categories multiplied, we would be able to see what gendered characteristics were most salient to the people writing during a specific time period, finding *their* gender cues rather than always *our own*. The M/F opposition is implicitly and always aggressive in the sense of promoting oneself at the expense of others, sometimes others who lived in the past.³⁰ Computer screens as opposed to stereotyped pages easily afford the fluid exploration of parameters and taxonomies, through which many sorts of experiments can be tested: interactive visualizations can give us not objective answers rooted in aggressively reductive oppositions, but parallax, multiple perspectives for viewing a very complex reality. In analyzing gender, the computational analysis of culture should never be understood as an objective observer measuring an inert reality. Rather, it should be understood as a means of counting people counting. Gender analysis thus provides an important test case for cultural analytics, allowing the field to show how self-critical and historically reflexive it can be. But it does more.

Transforming the print humanities into the digital humanities offers us an opportunity.³¹ It is a very short step from “typically” to “stereotypically,” as short as it is from printed to stereotyped pages, which are notoriously difficult to tell apart (Shillingsburg). Thawing typographical fixity via computer screens, moving from private marginalia to public, scholarly, dynamic interaction, offers us the chance to contest discursive constructions like the sex/gender/sexuality system at every moment of their construction.

NOTES

1. This definition of “cultural analytics” comes from *CA: The Journal of Cultural Analytics* (<http://culturalanalytics.org/about/>). The term’s origin and scope are defined in note 7. I use the term because (1) I think it is an emerging field or subfield in Literary and Cultural Studies, and (2) the emphasis on “culture” opens the door to cultural theory,

a most welcome addition to data-mining techniques (see Klein, “Distant Reading after Moretti”).

2. Rule #1 in Catherine D’Ignazio and Lauren Klein’s essay, “Feminist Data Visualization,” is to “Rethink Binaries” (sec. 3.1).

3. Toril Moi (65). Whereas Deborah Tannen is very careful to use the terms “women” and “girls,” “men” and “boys,” throughout her work and deplors any essentialist interpretation of the notion of gender difference in conversational styles formulated by researchers like herself who take an interactional sociolinguistic approach (8–9, 12–13), “gender” and “biological sex” are sometimes used interchangeably in the introduction to Harrington et al., eds., *Gender and Language Research Methodologies* (2008). A large number of sociolinguists are now investigating “the intersection of gender, sex, and sexuality” (Levon, 300), which is to say separating them from each other and then investigating how these categories “mutually constitute” each other (Levon, 298, emphasis in the original; Levon offers full citations of the sociolinguists doing this work).

4. Stoller’s original work in a 1964 essay opens by distinguishing “gender identity” from “sex” (M/F), but later posits three forces: (1) culture/psychology, (2) anatomy, and (3) a “biological force” (220). In one of the two cases he presents, the “biological force” is congruent with anatomy: a child raised as a girl who wanted to be a man turns out to have male chromosomes (part of the “biological force”) and a very small penis (“anatomy”); in the second one, the famous case of Agnes, genitalia are male but the female biological force propels Agnes forward as female. Stoller recanted in a 1968 article, repudiating his own notion of “biological force” after discovering that Agnes had been secretly taking her mother’s hormone replacement medicines (83–84).

5. See <https://artfl-project.uchicago.edu/>.

6. Originally a printing term, the word “stereotype” later acquired negative connotations. Dictionaries typically cite Walter Lippmann’s 1922 essay, “Public Opinion,” as the origin of a derogatory definition of the word “stereotype,” but Walter Johnson notes that, before emancipation in the United States, the same stereotype plate depicting a slave was used over and over again in runaway slave advertisements (189), indicating racist origins for the term’s use. I am grateful to Lauren Klein for pointing this out.

7. The broader term that spans not simply literature but also all cultural artifacts, “cultural analytics,” was coined by Lev Manovich in 2007–2009 (“Cultural Analytics: Analysis and Visualization”; “Cultural Analytics”). Discussions about the value of computational work in literature often call it “distant reading,” first coined in 2000 by Franco Moretti in “Conjectures on World Literature,” who debated its value with Katie Trumpener in the pages of *Critical Inquiry* (see Moretti, “Style, Inc.,” and “Relatively Blunt”; Trumpener, “Critical Response 1”; see also Moretti, *Distant Reading*). For more recent discussions, see Bode (“Equivalence”) and Underwood (“Genealogy”).

8. Critiques of Watt’s masculinist argument about the English novel began in 1986 with Dale Spender in *Mothers of the Novel* (115–18, 138–39), to which Ian Watt responded (“Spencer on Trial”). This account comes from Josephine Donovan (441–42). For opposing views of Watt’s thesis on the French novel, see Miller.

9. For data mining that traces the “volatility” of gender categories, see Underwood, “Instability,” and chapter 4 of *Distant Horizons*, “The Metamorphoses of Gender.”

10. Chawton House, <https://chawtonhouse.org/>.

11. While this chapter was under submission, an excellent essay about how to avoid gender bias in Natural Language Processing, including stylometry, was published, which describes in detail how stereotypes and bias inform the process and results. See Koolen and van Cranenburgh; thanks to Andrew Piper for bringing this article to my attention.

12. On reproducibility in data mining, see Goldstone (“Reproducible”; thanks to Lauren Klein for this reference). Failed attempts to reproduce results among Digital Humanists, some even using the same data and algorithms on different machines, prompted Alan Liu to call for scientific workflow management (“Assessing Data Workflows”).

13. I am thinking here of the work done by Mary Douglas in *Purity and Danger*. The violence elicited when gender attributes are “out of place” (40) suggests that, as Douglas argues, gender is serving as a symbol for the cultural system as a whole (3–4, 114). As virulent reactions to gender ambiguity suggest, undermining that specific symbolic distinction poses a threat to the whole system that people have used to sort and interpret experience throughout their lives (121).

14. Coe’s account of the failure to measure effect size in educational research because of its relative neglect in the field of statistics is worth quoting in full: “The routine use of effect sizes, however, has generally been limited to meta-analysis—for combining and comparing estimates from different studies—and is all too rare in original reports of educational research (Keselman et al., 1998). This is despite the fact that measures of effect size have been available for at least sixty years (Huberty, 2002), and the American Psychological Association has been officially encouraging authors to report effect sizes since 1994—but with limited success (Wilkinson et al., 1999). Formulae for the calculation of effect sizes do not appear in most statistics textbooks (other than those devoted to meta-analysis), are not featured in many statistics computer packages, and are seldom taught in standard research methods courses. For these reasons, even the researcher who is convinced by the wisdom of using measures of effect size, and is not afraid to confront the orthodoxy of conventional practice, may find that it is quite hard to know exactly how to do so” (1).

15. For me, Hyde’s meta-analysis of M/F studies severely undermines biological determinism. In any case, in hermeneutics as opposed to science, “it’s natural” prevents further analysis. (As is well known, Edmund Burke began deploying “the natural” as an ideological tool in his pamphlet condemning the French Revolution published in 1789, *Reflections on the Revolution in France*). Analyzing society at all requires dislodging biologicistic conceptions. For example, for Erving Goffman, the practices of con men reveal the cues we use to present ourselves as something to someone else, however unconsciously, and drag can do the same for gender (Butler). Tallying up those cues and historicizing them promote analytic thinking.

16. See Mandell, “Gendering Digital Literary History.”

17. The original figure appears here, <http://www.dh2012.uni-hamburg.de/wp-content/uploads/2012/07/img154-3.jpg>, and the revised figure here: <http://www.matthewjockers.net/macroanalysisbook/color-versions-of-figures-9-3-and-9-4/gendernet/>.

18. Ryan and Heuser (46, qtd. in Liu, “What Is the Meaning,” 411; emphasis added by Liu).

19. The novels he used, except for the canonical ones he added, are listed here: <https://chawtonhouse.org/the-library/library-collections/womens-writing-in-english/novels-online/>. Behn’s *Oroonoko*, Lennox’s *The Female Quixote*, Charke’s *The Memoirs of Charlotte Charke*, and Owenson’s *The Missionary*, not collected by the Chawton-House donor, contain highly adventurous content. Sentimentality would not be seen as distinctively “F” if we added Mackenzie’s *Man of Feeling*, Fielding’s *Shamela*, and all of Godwin’s novels. Less certain, though worth testing, is that “breast” might not be associated with “F” if we added Cleland’s *Fanny Hill* or translations of Rousseau’s *Julie, ou la Nouvelle Heloise*, or Diderot’s *Bijoux Indiscret* or *La Religieuse*.

20. Obviously, one of these tasks is viewed as a “natural” process and the other as an intellectual process. In my view, this difference results from how early in human development the intellectual task of acquiring the ability to perform gender occurs. But it helps to remember that learning gender cues is coeval with another intellectual task, linguistic acquisition, and that speaking too ultimately feels like a natural capacity.

21. Thanks to Susan Brown for this point.

22. I presented this list at Dartmouth (2016) and Hamburg (2016: at 35.18 in the video stream), and it has subsequently been summarized and described in Drucker (“Why Distant Reading Isn’t”). She concludes, “The point is that the model [of gender categories] determines what will emerge from analysis of the texts” (632).

23. This portion of Underwood’s blog posting is very tongue in cheek, as this comment about being a “pair of gametes” indicates: Underwood is irritated, as he goes on to show, by having had to court complexity in his analysis of genre in order to disarm his critics. But however sarcastically intoned, the point is worth pursuing: a pair of gametes becomes very complex indeed, and all the zeros and ones at the lowest level of machine language can similarly generate what may be an unfathomable complexity.

24. See <http://www.voyant-tools.org/>.

25. See <http://methodi.ca/>.

26. See Andrew Piper’s introduction to the new section of this journal: <http://culturalanalytics.org/2017/10/introducing-data-sets-a-new-section/>.

27. This work is currently being undertaken by the 3DH group, for whom the third dimension is not spatial but hermeneutic: <http://threedh.net/3dh/>; see Drucker (“Non-Representational Approaches”) for a full explanation of the project’s achievements so far and proposed future developments. The project has received additional funding as the for-Text project: <http://fortext.net/>.

28. Parody may be more useful for assimilating norms than it is for debunking them. The most amazing instance of this that I know of is the fact that early eighteenth-century political satirists wrote repeatedly of King George’s “prime minister” as a way of denigrating Sir Robert Walpole: they felt that none of the king’s ministers should be preferred. The satirists were laughing at Walpole for ingratiating himself and warning the populace that

one minister had too much power over state affairs. “Prime minister” is of course now a title in England of a very powerful person indeed (Goldgar).

29. On the renewed importance of “modesty” in literary criticism, see Jeffrey Williams.

30. Roland Barthes’s 1977 essay “The Death of the Author” foregrounds the belligerence of immodest literary criticism: “victory to the critic” (147).

31. By “print humanities,” I mean the methods of humanistic study that originated in mass printing. The transformation of modes of criticism is happening already, and the pitfalls of this process reside for me in failing to weed out the worst aspects of print culture and remediating digital media according to their requirements, fixity being one of them. See Mandell, *Breaking the Book*.

BIBLIOGRAPHY

- Barthes, Roland. “The Death of the Author.” In *Image Music Text*, 142–48. 2nd ed. New York: Hill and Wang, 1994.
- Bode, Katherine. “The Equivalence of ‘Close’ and ‘Distant’ Reading: Or, toward a New Object for Data-Rich Literary History.” *MLQ* 78, no. 1 (2017): 77–104.
- Burrows, James. *Computation into Criticism: A Study of Jane Austen’s Novels and an Experiment in Method*. Oxford: Clarendon Press, 1987.
- Butler, Judith. *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge, 1990.
- Cartwright, Nancy L., and Norman M. Bradburn. “A Theory of Measurement.” Working paper. London: London School of Economics, 2010, <http://dro.dur.ac.uk/18310/1/18310.pdf>. Accessed July 9, 2016.
- Coe, Robert. “It’s the Effect Size, Stupid. What Effect Size Is and Why It Is Important.” Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, England, September 12–14, 2002, <http://www.leeds.ac.uk/educol/documents/00002182.htm>. Accessed December 31, 2015.
- Cole, Elizabeth R. “Intersectionality and Research in Psychology.” *American Psychologist* 64, no. 3 (2009): 170–80.
- D’Ignazio, Catherine, and Lauren F. Klein. “Feminist Data Visualization.” Workshop on Visualization for the Digital Humanities. IEEE VIS 2016, Baltimore, Md., 2016, http://www.kanarinka.com/wp-content/uploads/2015/07/IEEE_Feminist_Data_Visualization.pdf.
- Donovan, Josephine. “Women and the Rise of the Novel: A Feminist-Marxist Theory.” *Signs* 16, no. 3 (1991): 441–62.
- Douglas, Mary. *Purity and Danger: An Analysis of the Concepts of Pollution and Taboo*. New York: ARK, 1966, 1988.
- Drucker, Johanna. *Graphesis: Visual Forms of Knowledge Production*. Cambridge, Mass.: Harvard University Press, 2014.
- Drucker, Johanna. “Non-Representational Approaches to Modeling Interpretation in a Graphical Environment.” *Digital Scholarship in the Humanities* 33, no. 2 (2018): 248–63, <https://doi.org/10.1093/lc/fqx034>.
- Drucker, Johanna. “Why Distant Reading Isn’t.” *PMLA* 132, no. 3 (2017): 628–35.

- Fitz-Gibbon, C. T. "Meta-Analysis: An Explication." *British Educational Research Journal* 10, no. 2 (1984): 135–44.
- Flanders, Julia. "Detailism, Digital Texts, and the Problem of Pedantry." *TEXT Technology* 2 (2005): 41–70.
- Forrester, John. *Thinking in Cases*. Malden, Mass.: Polity, 2017.
- Gelfand, Elissa, and Margaret Switten. "Gender and the Rise of the Novel." *French Review* 61, no. 3 (1988): 443–53.
- Goffman, Erving. *The Presentation of Self in Everyday Life*. New York: Anchor Doubleday, 1959. First published 1956.
- Goldgar, Bertrand G. *Walpole and the Wits: The Relation of Literature and Politics, 1722–1742*. Lincoln: University of Nebraska Press, 1976.
- Goldstone, Andrew. "From Reproducible to Productive." *CA: Journal of Cultural Analytics*, February 27, 2017, <http://culturalanalytics.org/2017/02/from-reproducible-to-productive/>. Accessed December 15, 2017.
- Gorelick, Sherry. "Contradictions of Feminist Methodology." *Gender and Society* 5, no. 4 (1991): 459–77.
- Haraway, Donna J. "Gender for a Marxist Dictionary: The Sexual Politics of a Word." In *Simians, Cyborgs, and Women: The Reinvention of Nature*. London: Free Association Books, 1991. Rpt. in *Women, Gender, Religion: A Reader*, edited by Elizabeth A. Castelli, 49–75. New York: Palgrave, 2001.
- Haraway, Donna J. *Modest_Witness@Second_Millennium.FemaleMan©_Meets_Onco Mouse™*. New York: Routledge, 1997.
- Hardt, Moritz. "How Big Data Is Unfair." *Medium*, September 26, 2014, <https://medium.com/@mrtz/how-big-data-is-unfair-9aa544d739de#.hmqcrs7sr>. Accessed December 10, 2015.
- Harrington, Kate. "Perpetuating Difference? Corpus Linguistics and the Gendering of Reported Dialogue." In *Gender and Language Research Methodologies*, edited by Kate Harrington, Lia Litosseliti, Helen Sauntson, and Jane Sunderland, 85–102. New York: Palgrave Macmillan, 2008.
- Hartsock, Nancy C. M. *Money, Sex, and Power: Toward a Feminist Historical Materialism*. New York: Longman, 1985.
- Hesse, Mary. "Review, *The Structure of Scientific Revolutions*." *Isis* 54, no. 2 (1963): 286–87.
- Heuser, Ryan, and Long Le-Khac. "Learning to Read Data: Bringing Out the Humanistic in the Digital Humanities." *Victorian Novels* 54, no. 1 (2011): 79–86.
- Hodgson, Thomas. *An Essay on Stereotype Printing, Including a Description of the Various Processes*. Newcastle, England: Hodgson, 1820.
- Huberty, C. J. "A History of Effect Size Indices." *Educational and Psychological Measurement* 62, no. 2 (2002): 227–40. [Cited by Coe.]
- Hyde, Janet Sibley. "The Gender Similarities Hypothesis." *American Psychologist* 60, no. 6 (2005): 581–92.
- Jockers, Matthew L. "Computing and Visualizing the 19th-Century Literary Genome." *DH 2012 University of Hamburg Conference Proceedings*, July 16–22, 2012, <http://www>

- .dh2012.uni-hamburg.de/conference/programme/abstracts/computing-and-visualizing-the-19th-century-literary-genome.1.html. Accessed March 21, 2017.
- Jockers, Matthew L. "The LDA Buffet Is Now Open: Or, Latent Dirichlet Analysis for English Majors." September 29, 2011, <http://www.matthewjockers.net/2011/09/29/the-lda-buffet-is-now-open-or-latent-dirichlet-allocation-for-english-majors/>. Accessed April 19, 2016.
- Jockers, Matthew L. *Macroanalysis: Digital Methods and Literary History*. Chicago: University of Illinois Press, 2013.
- Jockers, Matthew L., and Julia Flanders. "A Matter of Scale." DigitalCommons@University of Nebraska—Lincoln. March 18, 2013, <http://digitalcommons.unl.edu/englishfacpubs/106/>. Accessed April 19, 2016.
- Jockers, Matthew L., and Gabi Kirilloff. "Understanding Gender and Character Agency in the 19th Century Novel." *CA: Journal of Cultural Analytics*, December 1, 2016, <http://culturalanalytics.org/2016/12/understanding-gender-and-character-agency-in-the-19th-century-novel/>. Accessed March 18, 2017.
- Johnson, Walter. *Soul by Soul: Life inside the Antebellum Slave Market*. Cambridge, Mass.: Harvard University Press, 2009.
- Keselman, H. J., C. J. Huberty, L. M. Lix, S. Olejnik, et al. "Statistical Practices of Educational Researchers: An Analysis of Their ANOVA, MANOVA, and ANCOVA Analyses." *Review of Educational Research* 68, no. 3 (1998): 350–86. [Cited by Coe.]
- Klein, Lauren. "Distant Reading after Moretti." MLA 2018, <http://lklein.com/2018/01/distant-reading-after-moretti/>. Accessed March 13, 2018.
- Koolen, Corina, and Andreas van Cranenburgh. "These Are Not the Stereotypes You Were Looking For: Bias and Fairness in Authorial Gender Attribution." In *Proceedings of the First Workshop on Ethics in Natural Language Processing*, edited by Dirk Hovy, et al., 12–22/ Stroudsburg, PA: Association for Computational Linguistics (ACL): 2017.
- LaBrada, Eloy. "Categories We Die For: Ameliorating Gender in Analytic Feminist Philosophy." *PMLA* 131, no. 2 (2016): 449–59.
- Levon, Erez. "Integrating Intersectionality in Language, Gender, and Sexuality Research." *Language and Linguistics Compass* 9, no. 7 (2015): 295–308.
- Liu, Alan. "Assessing Data Workflows for Common Data 'Moves' across Disciplines." Alan Liu (personal blog). May 6, 2017, <http://liu.english.ucsb.edu/data-moves>. Accessed December 15, 2017. doi: 10.21972/G21593.
- Liu, Alan. "What Is the Meaning of the Digital Humanities to the Humanities?" *PMLA* 128, no. 2 (2013): 409–23.
- Maccoby, Eleanor and Carol Jacklin. *The Psychology of Sex Differences*. Stanford, Calif.: Stanford University Press, 1974.
- Mandell, Laura. *Breaking the Book: Print Humanities in the Digital Age*. Malden, Mass.: Wiley Blackwell, 2015.
- Mandell, Laura. "Gender and Big Data: Finding or Making Stereotypes?" Presentation at Dartmouth College, Hanover, N.H., May 24, 2016.

- Mandell, Laura. "Gendering Digital Literary History: What Counts for Digital Humanities." In *A New Companion to Digital Humanities*, edited by Susan Schreibman, John Unsworth, and Ray Siemens, 511–24. 2nd ed. Malden, Mass: Wiley-Blackwell, 2016.
- Mandell, Laura. "Visualizing Gender Complexity." Universität Hamburg. June 9, 2016, <http://lecture2go.uni-hamburg.de/l2go/-/get/v/19498>.
- Manovich, Lev. "Cultural Analytics: Analysis and Visualization of Large Cultural Data Sets." 2007–2008, http://softwarestudies.com/cultural_analytics/cultural_analytics_2008.doc. Accessed March 13, 2018.
- Manovich, Lev. "Cultural Analytics: Visualizing Cultural Patterns in the Age of 'More Media.'" 2009, http://manovich.net/content/04-projects/063-cultural-analytics-visualizing-cultural-patterns/60_article_2009.pdf [The article did not appear in DOMUS, as was planned.]
- Marcus, Sharon, Heather Love, and Stephen Best. "Building a Better Description." *Representations* 135 (Summer 2016): 1–21.
- Merriman, Ben. "A Science of Literature: Review of *Distant Reading* (Moretti), *The Bourgeois* (Moretti), and *Macroanalysis* (Jockers)." *Boston Review*, August 3, 2015, <http://bostonreview.net/books-ideas/ben-merriman-moretti-jockers-digital-humanities>. Accessed January 7, 2017.
- Miller, Nancy K. "Men's Reading, Women's Writing: Gender and the Rise of the Novel." *Yale French Studies* 75 (1988): 40–55.
- Moi, Toril. *Sexual/Textual Politics: Feminist Literary Theory*. New York: Methuen, 1985.
- Moretti, Franco. "Conjectures on World Literature." *New Left Review* 1 (2000): 54–66.
- Moretti, Franco. *Distant Reading*. London: Verso, 2013.
- Moretti, Franco. "'Operationalizing': Or, the Function of Measurement in Modern Literary Theory." Stanford Literary Lab Pamphlet 6. December 2013, <http://litlab.stanford.edu/LiteraryLabPamphlet6.pdf>. Accessed April 20, 2016.
- Moretti, Franco. "Relatively Blunt." *Critical Inquiry* 36 (2009): 172–74.
- Moretti, Franco. "Style, Inc. Reflections on Seven Thousand Titles (British Novels, 1740–1850)." *Critical Inquiry* 36 (2009): 134–58.
- Noble, Safiya Umoja. "Google Search: Hyper-Visibility as a Means of Rendering Black Women and Girls Invisible." *In Visible Culture* 19 (Fall 2013).
- Nuwer, Rachel. "Data Mining the Classics Clusters Women Authors Together, Puts Melville out on a Raft." *Smithsonian Smart News*, August 27, 2012, <http://www.smithsonianmag.com/smart-news/data-mining-the-classics-clusters-women-authors-together-puts-melville-out-on-a-raft-16028354/>. Accessed March 23, 2017.
- Olsen, Mark. "Écriture feminine: Searching for an Indefinable Practice?" *Literary and Linguistic Computing* 20, suppl. (2005): 147–64.
- Olsen, Mark. "Qualitative Linguistics and *Histoire des Mentalités*: Gender Representation in the *Trésor de la Langue Française*." *QUALICO*, January 14, 1992.
- O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown, 2016.

- Pennebaker, James. *The Secret Life of Pronouns*. New York: Bloomsbury USA, 2011.
- Richardson, Sarah. *Sex Itself: The Search for Male and Female in the Human Genome*. Chicago: University of Chicago Press, 2013.
- Rockwell, Geoffrey, and Stéfan Sinclair. Method Commons. <http://www.methodi.ca.org>. Accessed February 12, 2019.
- Rockwell, Geoffrey, and Stéfan Sinclair. Voyant Tools. <http://www.voyant-tools.org>. Accessed February 12, 2019.
- Rybicki, Jan. "Vive la différence: Tracing the (Authorial) Gender Signal by Multivariate Analysis of Word Frequencies." *Digital Scholarship in the Humanities* (2015): 1–16. doi: 10.1093/lhc/fqv023.
- Scott, Joan W. "Gender: A Useful Category of Historical Analysis." *American Historical Review* 91, no. 5 (1986): 1053–75.
- Shillingsburg, Peter L. "Detecting the Use of Stereotype Plates." *Editorial Quarterly* 1, no. 1 (1975).
- Spender, Dale. *Mothers of the Novel: 100 Good Women Writers before Jane Austen*. New York: Pandora, 1986.
- Sprague, Joey, and Mary Zimmerman. "Quality and Quantity: Reconstructing Feminist Methodology." *American Sociologist* 20, no. 1 (1989): 71–86.
- Stengers, Isabelle. *Cosmopolitics I*. Translated by Robert Bononno. Minneapolis: University of Minnesota Press, 2010.
- Stoller, Robert. "A Contribution to the Study of Gender Identity." *International Journal of Psychoanalysis* 45, nos. 2–3 (1964): 220–26.
- Stoller, Robert. "A Further Contribution to the Study of Gender Identity." *International Journal of Psychoanalysis* 49, nos. 2–3 (1968): 364–69.
- Tannen, Deborah. *Gender and Discourse*. New York: Oxford University Press, 1994.
- Trumpener, Katie. "Critical Response I: Paratext and Genre System." *Critical Inquiry* 36 (2009): 159–70.
- Underwood, Ted. *Distant Horizons: Digital Evidence and Literary Change*. Chicago: University of Chicago Press, forthcoming.
- Underwood, Ted. "A Genealogy of Distant Reading." *DHQ* 11, no. 2 (2017), <http://www.digitalhumanities.org/dhq/vol/11/2/000317/000317.html>. Accessed December 15, 2017.
- Underwood, Ted. "Measurement and Modeling." The Stone and the Shell (personal blog). December 20, 2013, <https://tedunderwood.com/2013/12/20/measurement-and-modeling/>. Accessed April 20, 2016.
- Underwood, Ted. "The Real Problem with Distant Reading." The Stone and The Shell (personal blog). May 29, 2016, <https://tedunderwood.com/2016/05/29/the-real-problem-with-distant-reading/>. Accessed December 15, 2017.
- Underwood, Ted, and David Bamman. "The Instability of Gender." The Stone and the Shell (personal blog). January 9, 2016, <http://tedunderwood.com/2016/01/09/the-instability-of-gender/>. Accessed April 19, 2016.

- Unsworth, John. "Documenting the Reinvention of Text: The Importance of Failure." *Journal of Electronic Publishing* 3, no. 2 (1997), <https://quod.lib.umich.edu/jjep/3336451.0003.201/--documenting-the-reinvention-of-text-the-importance?rgn=main;view=fulltext>. Accessed April 19, 2016.
- Watt, Ian. "Spender on Trial." *The World and I*. November 1987, 351–56.
- Westbrook, Laurel, and Kristen Schilt. "Doing Gender, Determining Gender: Transgender People, Gender Panics, and Maintenance of the Sex/Gender/Sexuality System." *Gender & Society* 28, no. 1 (2014): 32–57.
- Wilkinson, L., and Task Force on Statistical Inference, APA Board of Scientific Affairs. "Statistical Methods in Psychology Journals: Guidelines and Explanations." *American Psychologist* 54, no. 8 (1999): 594–604. [Cited by Coe.]
- Williams, Jeffrey J. "The New Modesty in Literary Criticism." *Chronicle of Higher Education* 61, no. 17 (2015): B6–B9.
- [Wollstonecraft, Mary.] *A Vindication of the Rights of Men, in a Letter to the Right Honourable Edmund Burke*. London: J. Johnson, 1790.
- Wollstonecraft, Mary. *A Vindication of the Rights of Woman*, 3rd ed., 1792. Edited by Deidre Shauna Lynch. New York: W. W. Norton, 2009.