# Bringing Computation into Cultural Theory: Four Good Reasons (and One Bad One)

Clayton Childress

We used to talk. By "we" I mean cultural sociologists and scholars in the humanities, and by "used to talk" I mean acknowledge each other's existence, at times perhaps even generously so. There are different versions as to what happened, one of which is a bit more intellectual than the other, although neither of which are entirely right. The more intellectual version is that for a brief spell in the late 1980s and early 1990s it looked like our interests might converge. At around the same time, many of us stopped being scolds about popular culture, deciding instead that it was more fruitful and interesting to engage the world than to police it. Some of us were also asking similar questions, be it about the role of authors and their ability (or lack thereof) to enforce, guide, or push readers into certain meanings, or about the role of interpretive communities and groups either to buffer against the impingement of those meanings or to generate localized meanings all anew. So we congregated around folks such as I. A. Richards, Wolfgang Iser, Hans Robert Jauss, Mikhail Bakhtin, Stanley Fish, Roland Barthes, or Michel Foucault, and sometimes we even cited each other too, and then it just all kind of petered out.

One reason for that petering out, at least within the discipline of sociology, was that the production of culture approach was ascendant and reoriented questions about authorial creation into questions about institutional production. At the same time, studies of cultural reception could only "discover" groups-based heterogeneity in meaning making so many times over before there wasn't anywhere else to go. The Frankfurt School had, for us at least, proven to be an empirical dead end, and now, fully out in the world studying actual people and the meanings they made of texts, we got deeply skittish about our own interpretive meanings, which became a bridge that we as cultural sociologists could no longer pass.

The second reason for why we first started talking and then stopped is less intellectual but perhaps truer: we were only ever talking because people *made that talking happen.* And then the next cohort came along,

*New Literary History,* 2022 & 2023, 53 & 54: 975–983

and, as is usually the case, they were quite busy making different things happen, and the connection was lost as connections are usually lost, which is in silence. Wendy Griswold has been for me a (if not the) central figure of that earlier generation of integrative talkers. Griswold, who had graduate training in both English and sociology, and who was Bergen Evans Professor in the Humanities at Northwestern until her retirement this year, makes clear the benefits of our cross pollination. Her blending of the humanistic with the sociological—be it about the "complex" of Nigerian literature, "cultural power" and the reception of George Lamming novels across international contexts, or the role of copyright law in the divergence and then convergence of topics in US and UK literatures—was a clarion call for me that this type of multivocal work not only could be done but should be done.[1] And in 1992, noticing an institutionalization of the sociology of culture, Griswold took the occasion to write a brief essay about four good debates (and one bad one) in that process.[2] The title and format of this essay owes a direct debt to hers. As we're beginning to talk again, and this time through the incorporation of computation into cultural theory, what are good and bad reasons to do so? In what follows I make some claims, which as usual, might be wrong.

## Good Reason #1: Reconnection and New Ideas

As already discussed, the integration of computation into cultural theory has provided an occasion for scholars from cultural sociology and the humanities to start talking again, and that alone is reason to celebrate. A key insight of sociology is that oftentimes communities matter much more to people than the ostensible "reason" for their existence does. It's why it's not hard to find nonbelievers in church, why people who don't much like the sun can still be found at the beach, and why, without having met you, I'd probably be right in guessing that if you live in the US you've both been to a baseball game and don't much care about baseball. This is Émile Durkheim's classic insight about groups' celebration of themselves, their groupness, and the joys of belonging to something together.[3] Through this logic, that the integration of computation into cultural theory has us talking again should be reason enough to foster interaction around the topic.

Beyond the reasons of community, there are also instrumental reasons for us to talk again, as integration across groups is where good ideas come from.[4] Be they groups of different discursive styles, different sociodemographic backgrounds, or groups that incorporate different

disciplinary combinations of ideas, it's in these types of groups where the most useful and generative ideas form.[5] And there's evidence that this type of generative boundary spanning is already happening. With some peers in English, I can now communicate quite fluently about bag-of-words versus embedding models, whereas with others I can also communicate quite fluently about the bag of words that is the writing of Pierre Bourdieu. And, without any obsequiousness at all, I not only admire Richard Jean So's *Redlining Culture* and Ted Underwood's *Distant Horizons*, but also first learned about them from other cultural sociologists who were raving about them too (Jennifer Lena and Paul DiMaggio, respectively).[6] That should tell you something. And even if sometimes things don't run as smoothly—when, for instance, scholars in the humanities think of my efforts in *Under the Cover* to quantify and test authorial intentionality as a bit goofy, anachronistic, or missing the point (and from what I've heard, some do)—that's okay too.[7] We don't have to agree all the time, nor should we; otherwise, we wouldn't be benefiting from our cross pollinations and generating new (and sometimes bad) ideas through stretching across our differences.

## Good Reason #2: New Vantage Points into Power and Inequality

One of the critiques of the sociology of culture, particularly as it related to media industries, was that it lacked a focus on power and inequality.[8] For reasons that were not entirely unfounded, there was some truth to this, but not anymore. In fact, one way our interests are starting to converge with those of our friends in the humanities is precisely on these topics. Most centrally, this shared interest is still operating at the micro- and meso-levels, making our efforts not only more amenable to integrating computation into cultural theory, but also more engaged with problems as targetable and fixable at the levels of careers, organizations, and industries. In sociology, much of this new work interrogates the way that identity categories intersect with other valuation, evaluation, and classificatory systems, leading to systemic-level racism, sexism, and classism through the repeated glancing cuts of unequal sorting into categories rather than into total exclusion.[9]

For instance, in my own work, I use literary agents' self-reported areas of interest to show that cultural matching happens along sociodemographic lines, such that men agents are less likely than women agents to represent women's fiction, older agents are less likely than younger agents to represent young adult fiction, and white agents are less likely

than are non-white agents to represent ethnic and multicultural fiction.[10] This type of sorting based on the agents' own tastes, interests, and experiences would be non-problematic if agents were themselves sociodemographically diverse, which they are not. Back in the late 2010s when I was collecting my data, the plurality of agents (about 40%) were white women working in New York City, and I could identify fewer than five Black men simultaneously working as literary agents in the US. Laura B. McGrath's research on "comping white" similarly shows how in isolation from each other processes of comparison and classification can appear obvious and banal, while in the aggregate and at the systemic level these decisions reveal a literary prototype—both for type and for improbable but optimistic sales expectation—that is overwhelmingly white.[11] While industry-based data initiatives like the Lee & Low Diversity Baseline Survey, the VIDA Count, and *Publisher's Weekly* Publishing Industry Salary Survey have ameliorated some of our neglect at really getting our hands dirty with data and inequality in the literature, we have more to offer than we've given so far. This is the type of work we can, and should, do together.

## Good Reason #3: Getting Off the Main Effect

Oftentimes, both our methods and our theories focus on the main effect. By this I mean we imprint monocausal models onto a multicausal world. We do so for the sake of sometimes necessary simplification—a truly accurate map of the world would be the size of the world, and therefore useless as a map—but we obscure both multicausal phenomena and the underlying heterogeneity in what we claim to be explaining when we do. Computational methods are generally good at capturing this type of usually obscured heterogeneity, be it for how the inputs of racial classifications vary for different raced groups, or for how the relationship between unusual genre combinations and popularity in music follows a general pattern that varies based on geographic locale.[12] At a higher level of abstraction, while individuals and the institutions they make real generally seem to trend toward reification and reproduction, both individuals and institutions also fall off these trajectories quite frequently. A hybridization of methods and theories that can account for both the main effect (of reproduction) and variations on that effect (that leads to change) is therefore necessary.

Getting off the main effect and more oriented toward the heterogeneity within that main effect may be particularly useful for the types of classificatory and categorical thinking that some cultural sociologists

and humanists share. Quite frequently and without even really thinking about it, we turn the world into dummy variables: a landscape of conversational "1"s and "0"s. While this is useful for the purposes of mobilization—political activism, motivational speeches in sports or life—these simplifications may sometimes be less useful when trying to explain the world. For example, from computational methods we know that, be it literature, painting, or music, styles and types seem to slowly build and then fall rather than outright replace each other in a clearcut binary.[13] From the methodological ability to see these ebbs and flows our theories can improve too, allowing us to better capture, make sense of, and convey a world of continuums, graded memberships within and across categories, fuzzy sets, probabilistic (rather than definitive) assignment, and so on, leading to generally overlooked heterogeneity *within* groups.[14] We can, I think, get farther away from theorizing around artificially clean and crisp categorical assignments that force us to studiously pay selective inattention to all those complicating shades of gray.[15]

## Good Reason #4: Inductive Methods and Theory Generation

Even though we know better, in sociology we tend to ascribe to a cognitively lazy shortcut in which qualitative research is all inductive (it surely isn't) and quantitative work is all deductive (it surely isn't either). We putter along as if this is just the way it works, and then on either side of this imaginary divide, we gussy up what we've got with the insertion of causal language on the back end.[16] This all works well enough once you've been socialized into the rules of the game, but to an outsider, it's bizarre. As Laura Nelson recently argued in a special issue of *Poetics* honoring the career of John Mohr (a pioneer of measuring meaning in sociology and my PhD advisor), machine learning breaks these assumptions, and excitingly so.[17] Machine learning, Nelson argues, is epistemologically aligned with theories of intersectionality in that they are both deductive approaches to the world designed to uncover unobserved combinatorial phenomena. She then shows evidence of this effect along lines of race and gender in first person narratives from the nineteenth-century US South.

As part of a burgeoning wave of "forensic" social science, Nelson, like Mohr before her, is giving us the language to more fully break our schematic associations in a qualitative/inductive and quantitative/deductive binary.[18] And while deductive approaches test theories, inductive approaches are more hospitable environs for *generating* theories and for

setting us off down whole new paths. Because of this, most promising, I think, is the role of machine learning in generating theories and in contributing to research that relies on methodological bricolage. This includes everything from pairing inductive interviews with deductive experiments when investigating resume "whitening" to integrating archival research with computation and close reading.[19] Multiple methods also facilitate multiple data sources—be it in the global diffusion and institutionalization of *One Hundred Years of Solitude* or placed-based path dependence in the second wave feminist movements—which move us beyond triangulation and "confirmatory" approaches and allow for "propulsive facilitation" across methods and data sources in which epistemic standpoints change as we go.[20] What this means is that computation not only illuminates our existing cultural theory but drives new cultural theory, popping us out of the occasional rut in which theories only make sense in relation to other theories, with the real world having been seemingly left behind at some point along the way.

## And One Bad Reason: Using Fancy Methods Because We Can

Too frequently, people use fancy methods the same way cops use Maglites: more for intimidation than for illumination. The friendlier version of this is using fancy methods for their own sake or for producing findings or aesthetically pleasing figures that don't tell us much beyond what we already knew. This can be found in the N-measuring contests of some big data research and in the cat hairball-style networks graphs that, while visually impressive, seem to suggest "everything is everything, man" on this ayahuasca trip of interconnectedness that we call life. This orientation can also be found in the urge to use machine learning when the good ol' punching bag of regular regression works just as well.[21]

Fortunately, advances in natural language processing and machine learning may be growing into a more mature stage. In their infancy, one could get by doing demonstration-style papers for these approaches, showing that complex methods were resonant with what we know, or think we know, about the world. For example, we can computationally find evidence for Julian Jaynes's theory of the bicameral mind in *The Iliad* and *The Odyssey*, just as we can computationally show that while the logics of class across the US in the twentieth century remained mostly consistent, markers of what constituted those logics changed.[22] These papers are both impressive and impressively important, but they double as introducing methods to cultural sociologists, which in some ways

takes weight off the importance of what the methods actually found. We are, I believe, at a turning point, where machine learning is now old enough to really have to prove its worth beyond the presumption that we should use fancy methods simply because we can. There will still be methodological improvements that require new demonstration papers, of course, but the ratio of demonstration-style papers to just paper-style papers that use newfangled methods has probably already inverted.[23]

The challenge here is that for most people across most disciplines these are black box methods, and that dramatically limits the range of innovation that's tolerated from them. As my friend Adam Slez likes to say, there are only three options with black box computation: it spits out what you already know, so it isn't necessary; it spits out something totally unfamiliar to what you already know, so something must be broken inside the box; or it presents a plausibly small variation or complication on what you knew and therefore is both "correct" and legitimate. That basically sounds like what we've already been doing even without these new methods. In the long run, rather than using new methods as totems through which to create the boundaries of a special guild (or even worse, a weapon to bash others), we should treat them as earthly and humanely flawed, mostly misused, and as frequently impractical as they actually are. Ultimately, if these methods and variations on them become everyday and commonplace affairs, and if they disappoint on their promises as much as everything else does, then we have won. And we've won doubly so, because they've brought us back into conversation with one another too.

University of British Columbia

## NOTES

1 Wendy Griswold, "The Fabrication of Meaning: Literary Interpretation in the United States, Great Britain, and the West Indies," *American Journal of Sociology* 92, no. 5 (1987): 1077–1117; Griswold, *Bearing Witness: Readers, Writers, and the Novel in Nigeria* (Princeton, NJ: Princeton Univ. Press, 2000); and Griswold, "American Character and the American Novel: An Expansion of Reflection Theory in the Sociology of Literature," *American Journal of Sociology* 86, no. 4 (1981): 740–65.

2 Griswold, "The Sociology of Culture: Four Good Arguments (and One Bad One)," *Acta Sociologica* 35, no. 4 (1992): 323–28.

3 Lyn Spillman, *Solidarity in Strategy: Making Business Meaningful in American Trade Associations* (Chicago: Univ. of Chicago Press, 2012).

4 Ronald S. Burt, "Structural Holes and Good Ideas," *American Journal of Sociology* 110, no. 2 (2004): 349–99.

5 Devah Pager, "Are Firms That Discriminate More Likely to Go Out of Business?," *Sociological Science* 3 (September 19, 2016): 849–59; Brian Uzzi et al., "Atypical Combinations and Scientific Impact," *Science* 342, no. 6157 (2013): 468–72; Matthew Corritore,

Amir Goldberg, and Sameer B. Srivastava, "Duality in Diversity: How Intrapersonal and Interpersonal Cultural Heterogeneity Relate to Firm Performance," *Administrative Science Quarterly* 65, no. 2 (2020): 359–94; and Yang Yang et al., "Gender-Diverse Teams Produce More Novel and Higher-Impact Scientific Ideas," *Proceedings of the National Academy of Sciences* 119, no. 36 (2022): e2200841119.

6 Richard Jean So, *Redlining Culture: A Data History of Racial Inequality and Postwar Fiction* (New York: Columbia Univ. Press, 2020); and Ted Underwood, *Distant Horizons: Digital Evidence and Literary Change* (Chicago: Univ. of Chicago Press, 2019).

7 Clayton Childress, *Under the Cover: The Creation, Production, and Reception of a Novel* (Princeton, NJ: Princeton Univ. Press, 2017).

8 David Hesmondhalgh, *The Cultural Industries* (SAGE Publications, 2002); Richard A. Peterson and N. Anand, "The Production of Culture Perspective," *Annual Review of Sociology* 30, no. 1 (2004): 311–34.

9 Maryann Erigha, "Racial Valuation: Cultural Gatekeepers, Race, Risk, and Institutional Expectations of Success and Failure," *Social Problems* 68, no. 2 (2021): 393–408; Alex van Venrooij, Candace Miller, and Vaughn Schmutz, "Race and Genre Ambiguity in the Critical Reception of Popular Music," *Sociological Inquiry* 92, no. 2 (2022): 568–96; and Anamik Saha and Sandra van Lente, "Diversity, Media and Racial Capitalism: A Case Study on Publishing," *Ethnic and Racial Studies* 45, no. 16 (2022): 216–36.

10 Childress and Jean-François Nault, "Encultured Biases: The Role of Products in Pathways to Inequality," *American Sociological Review* 84, no. 1 (2019): 115–41.

11 Laura B. McGrath, "Comping White," *Los Angeles Review of Books*, January 21, 2019, https://lareviewofbooks.org/article/comping-white/.

12 Ariela Schachter, René D. Flores, and Neda Maghbouleh, "Ancestry, Color, or Culture? How Whites Racially Classify Others in the U.S.," *American Journal of Sociology* 126, no. 5 (2021): 1220–63; and Daniel Silver et al., "Balancing Categorical Conventionality in Music," *American Journal of Sociology* 128, no. 1 (2022): 224–86.

13 Ke Nie, "Disperse and Preserve the Perverse: Computing How Hip-Hop Censorship Changed Popular Music Genres in China," *Poetics*, Measure Mohr Culture, 88 (October 1, 2021): 101590; Maximilian Schich et al., "A Network Framework of Cultural History," *Science* 345, no. 6196 (2014): 558–62; and Underwood, *Distant Horizons*.

14 Ellis P. Monk, "Inequality without Groups: Contemporary Theories of Categories, Intersectional Typicality, and the Disaggregation of Difference," *Sociological Theory* 40, no. 1 (2022): 3–27.

15 Eviatar Zerubavel, *The Fine Line: Making Distinctions in Everyday Life* (Chicago: Univ. of Chicago Press, 1993).

16 Gabriel Abend, Caitlin Petre, and Michael Sauder, "Styles of Causal Thought: An Empirical Investigation," *American Journal of Sociology* 119, no. 3 (2013): 602–54.

17 Laura K. Nelson, "Leveraging the Alignment between Machine Learning and Intersectionality: Using Word Embeddings to Measure Intersectional Experiences of the Nineteenth Century U.S. South," *Poetics*, Measure Mohr Culture, 88 (October 1, 2021): 101539.

18 Amir Goldberg, "In Defense of Forensic Social Science," *Big Data & Society* 2, no. 2 (December 1, 2015): 2053951715601145.

19 So, *Redlining Culture*; Sonia K. Kang et al., "Whitened Résumés: Race and Self-Presentation in the Labor Market," *Administrative Science Quarterly* 61, no. 3 (2016): 469–502.

20 Álvaro Santana-Acuña, *Ascent to Glory: How One Hundred Years of Solitude Was Written and Became a Global Classic* (New York: Columbia Univ. Press, 2020); Nelson, "Cycles of Conflict, a Century of Continuity: The Impact of Persistent Place-Based Political Logics on Social Movement Strategy," *American Journal of Sociology* 127, no. 1 (2021): 1–59; and Childress, Abigail Calonga, and Erik Schneiderhan, "Beyond Triangulation: Reconstructing

Mandela's Writing Life through Propulsive Facilitation at the Archive," *Qualitative Sociology* 43, no. 3 (2020): 367–84.

21 Matthew J. Salganik et al., "Measuring the Predictability of Life Outcomes with a Scientific Mass Collaboration," *Proceedings of the National Academy of Sciences* 117, no. 15 (2020): 8398–8403.

22 Dustin S. Stoltz and Marshall A. Taylor, "Concept Mover's Distance: Measuring Concept Engagement via Word Embeddings in Texts," *Journal of Computational Social Science* 2, no. 2 (2019): 293–313; Austin C. Kozlowski, Matt Taddy, and James A. Evans, "The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings," *American Sociological Review* 84, no. 5 (2019): 905–49.

23 On improvements and sensitivities see Ramina Sotoudeh and Paul DiMaggio, "Coping With Plenitude: A Computational Approach to Selecting the Right Algorithm," *Sociological Methods & Research*, September 13, 2021, 00491241211031273.