

## Coding, counting and cultural cartography

Monica Lee and John Levi Martin\*

Department of Sociology, University of Chicago, 1126 E 59th Street, Chicago, IL 60637, USA.

\*Corresponding author.

**Abstract** Richard Biernacki has, through painstaking attempts to replicate the codings of formal analyses of culture, concluded that such efforts to bring rigor into cultural analysis are futile. Coding does intrinsic violence to the nature of the material, and *imposes* interpretations as opposed to drawing them out in such a way that they can be made subject to critique. Here we argue that Biernacki's claims as to the intrinsic problems with coding are valid, but they do not necessarily imply that the only form of defensible analysis of cultural materials is a conventional humanistic one. Instead, they may just as well be taken to imply that we need to move *further* in the direction of formalism. Formal techniques that do not involve imposition of interpretation *before* the analysis, but rather condense information to facilitate an intersubjectively valid interpretation, do not suffer from the problems identified by Biernacki, and offer a path for a distinctly sociological contribution to cultural analysis. Further, although such techniques *simplify* their source works, they do so in the way a map simplifies – they make patterns accessible for joint exploration.

*American Journal of Cultural Sociology* (2015) 3, 1–33. doi:10.1057/ajcs.2014.13; published online 2 December 2014

**Keywords:** coding; interpretation; formal analysis; cultural networks

---

Certainly, every interpretation, if it wants to wring from what the words say what they want to say, must use violence. Such violence, however, cannot be roving arbitrariness.

Martin Heidegger, *Kant and the Problem of Metaphysics*

In *Reinventing Evidence in Social Inquiry*, Biernacki (2012a) makes a blistering critique of formal approaches to the examination of culture. Despite his close analysis (and relatively – perhaps surprisingly – positive reviews, Kuyvenhoven, 2013; Steinmetz, 2013; Riley, 2014), his work does not seem

to have led to any serious reflection among sociologists of culture, perhaps because of our tactful disattention to heated controversy, or – more likely – the tremendous evidentiary burden that would confront someone attempting to wade into the dispute (a key point to which we will return below). Further, the no-holds-barred approach of Biernacki's writing, and his tendency to treat large and small matters with equal weight, may have led sociologists of culture to assume that his core argument must be exaggerated or unfair.

We disagree, in that we think that the central arguments he has made are important and must be understood as positive, not negative, contributions to the sociology of culture. We understand that those whose work was subject to close critique did not necessarily feel favored by such treatment, but we see it as a testament to the role they have played in increasing the rigor of a sociological approach to culture. First, we should note that the reason why these examples were chosen by Biernacki is that they come from scholars who were willing to preserve and turn over their materials; further, Evans (2009a, b) had previously given close and detailed response to many of Biernacki's critiques. Many other scholars with lower standards would have discarded their notes (were any actually made), or at least claimed that they had, when faced with the prospect of a critical re-examination of their decisions.

And second, we pay tribute to Evans, Griswold, Stovel and Bearman, not simply because they and their work have served as inspirations for us, but because they were of sufficient caliber to serve as a foil for a serious thinking of fundamental issues. Few of our products could bear such weight. Still, Biernacki (2012a, p. 10) was quite correct – this is a new subfield in which 'norms are up for grabs', not in terms of intellectual honesty but in terms of what it means to competently marshal data to support a claim. Whether this is inherent in the formal analysis of culture as a 'blurred genre', or simply a function of its youth, however, we cannot say. To twist Weber's words, some subfields are cursed with eternal youth, and our goal is to see this one pushed safely into its middle age. We see the Biernacki exchange as the opportunity to determine whether norms can be developed in a way that leads to continued progress in the formal analysis of culture.

We note that Biernacki (2012a, p. 21) emphasized that the works he criticized 'positioned their methods in an ambiguous middle ground. None of the researchers resorted to reductive word or phrase searches, such as a computer executes. Instead, these sophisticated sociologists rejected cruder reduction and contrasted it to the more sensitive coding they sought to execute'. We will argue that this is exactly why his work can just as well support an argument for *increased* formalization as it can one for *decreased* formalization.

There are three levels of claims that Biernacki makes. The first is that coding cultural products obscures the interpretive moment; rather than facing up to the necessary challenge of the hermeneutic circle, coding (and choice of sampling frame) makes it a backstage *fait accompli*. The second is that coding goes further

and does violence to the materials, by reducing them to components that are re-assembled in a self-validating ritual. The third is the implication that we must return cultural products to the loving hands of the humanities, who will treat them with more respect.

We will go on to argue that all Biernacki's claims regarding the first point are correct, and we endorse them as a critique of coding. We do not focus on the ritualistic aspects of the second, because we recognize that there can be ritualistic aspects of scientific communication that do not undermine the claim of such communication to have a valid content. We accept Biernacki's second point – that coding does violence and permits arbitrary reassemblage – though we will attempt to at least highlight the problems for interpretation when analysts do *not* atomize. We will argue that Biernacki's third conclusion – back to Humanism! – does not logically follow from his first set of claims which rather, as hinted above, may justify more extensive formalization, but of one particular sort.

## Interpretation and Reinterpretation

### The fundamental problem

The core of the problem that confronts us is a familiar one, often called the hermeneutic circle, or the openness of interpretation. When we say what something (*X*) means, we invoke a second thing (*Y*). While in some contexts, such as translation, we go from one set to another (the meaning of *Messer* [German] is *knife* [English]), in the realm of cultural analysis, we are all in one very large set – a text itself, say, but even more, a family of texts. We can only stabilize our interpretation of some particular *X* by reference to our understanding of the whole network of meanings. For example, Melville described the whale in *Moby Dick* as having a pyramid-shaped hump on his back. Now if someone asks us, 'what does Moby Dick's hump mean?' we must appeal to Melville's writing in general; we may say 'it means a pyramid, but for Melville, such a pyramid itself means the radical indifference of nature to human concerns'. If someone doubts our interpretation, we may go outside this one text; perhaps we say 'look at how Melville uses the imagery of walls in "Bartleby" and then re-read this passage'. If our interlocutor is still doubtful we may go even further, and say 'you must compare this to other imagery of Egyptian and Assyrian ruins in 19th century English and American literature!' To understand what this one part of the whale means in the one book, we may need to understand nineteenth century English and American literature as a whole: yet we only understand this whole by understanding the parts.

Now this sort of complex regress does not always take place. If someone were to ask us 'what does Ahab being "dismasted" mean?' we might answer, 'it means

he is castrated'. And our interlocutor might simply grimace and say 'Ouch. I get it. No more needed'. But logically, it is interpretation all the way down, or all the way 'round and 'round.

In a quest to produce rigor and system in this messy world, the movement for the systematic approach to culture developed in sociology, especially in the 1980s and 1990s. This was, we argue, a salutary development, even though it does indeed suffer from the flaws identified by Biernacki. The idea was to take what sociologists are best at – constructing databases, establishing comparability and adjudicating hypotheses – and to transport it to the realm of culture. The scholars critiqued by Biernacki were each pathbreaking in this direction. The overall goal was a valuable one – to bring rigor to our methods, and to substitute intersubjectively valid and 'checkable' claims for personalistic ones.

As we will go on to argue, the way that this was done in the 1980s and 1990s *necessarily* involved 'coding' – the grouping of particularities into a category that was given a label. But as Biernacki has shown, this is equivalent to trying to solve the problem of the hermeneutic circle as Alexander the Great might – to slash it with a sword and stuff the pieces in a box out of sight. Rather than allow us to understand that the part is only interpreted by reference to the whole, the part is *fixed, determined*, stapled to its interpretation via the violence of 'proof by assertion'. Thus, the most obvious problem with coding is that it assumes what is to be proven. *Is this really a case of 'Occidentalism' (say)?* Simply *calling* it one, counting it thusly and burying the original, is hardly solving the problem of interpretation (also see Whewell, 1971 [1860]). Coding 'cannot take one beyond the confines of what has been generally supposed at the outset' (Biernacki, 2009a, p. 179). Sociologists tend to be nominalists, and therefore have a near infinite tolerance for allowing one another to define into existence the most doubtful of entities, and we think Biernacki's critique is crucial if only for this one point. But our argument is not that sociologists have botched their *execution* of coding; rather, it has to do with *inherent limitations* in the concept. Fleshing this out allows us to explain why we see coding as far from a minor issue of one particular approach.

Coding is an act of subsumptive judgment *par excellence*; that is, it is the predication of a particularity with a general concept by virtue of following a set of explicable rules. For example, Moby Dick, being of a species that has live birth and nursing, is a mammal. Such judgments can be *proven* to others via the use of concepts. But we are convinced by Kant (1987 [1790]) that in some circumstances – and we believe that cultural interpretation is one of them – we must also apply *reflective* judgments in which we attach a generality to a particularity without relying on explicit rules. This means that while we *can* achieve intersubjective concord with others about such judgments, we cannot, using concepts, *prove* our judgment to another who doubts us. This leads to a puzzle as to the scientific status of such judgments, and the hermeneutic tradition as explicated by Gadamer (1975 [1965]) begins from accepting the basic outlines of Kant's argument.

Thus, coding distills some of the fundamental difficulties in establishing a science of culture, but it simplifies and exaggerates them, making it a good case for thinking through our difficulties.<sup>1</sup> Coding is inherently a sort of interpretation that tends toward *totalism* (this text is a case of Occidentalism), one that tends to be *invisible* (in that the reader cannot see and critique the decisions), and one that leads to *nominalist perversities*, in which our main ‘conclusions’ come not from the pattern in the data, but how we happen to *label* the piles that we make out of it.

But the problem goes further and is more subtle, for as Biernacki (2009a, b) argues, the choice of the population from which to sample already greatly predetermines the conclusion. Before we can code documents to see whether a field has changed, say, we must first delimit (if not define) the field (on delimitation as opposed to definition, see Martin, 2014). In order to construct a population of X’s from which to sample, we need to know what an ‘X’ is. But we are often trying to determine the nature of X *through* our research, and so how we establish a sampling frame will affect, if not determine, our conclusions. It is all very well for Bourdieu to say (quite defensibly) that the field ends where the field effect ends, but the practicing scientist must make a decision, and an overwide net leads to overlong research programs.<sup>2</sup>

It gets even worse. As Biernacki (2012a; also 2014, p. 180) correctly says, when we construct codings, we often draw a rather generous line around our concept – but *illustrate* it using selective cases that are far from the median. In other words, we *work* in a nominalist fashion (‘by “hippies” I mean white Californians of no certain employment, between the ages of 16 and 28, who have long unkempt hair, beards, wear beads and walk barefoot’) but we *think* in a tendentiously ideal-typical fashion (‘for example, a classic hippy will be found in the case of Charles Manson ...’).

Now Biernacki (2009a, p. 121) also argues that coding fractures cultural works and allows the analyst to re-assemble the pieces in a new context of the analyst’s choosing. This may well be true, but it strikes us as very similar to what Latour (1999) argues is characteristic about *successful* science: that it involves transporting the subject to a site in which the analyst is at home and very possibly more powerful. Although we can recoil in prurient horror from Baconian and similar views of science as some violent ravishing or dismemberment of nature, it

---

<sup>1</sup> Further, to look for more ‘nuanced’ examples of coding, or to argue that there are ‘much better’ examples of coding in cultural sociology, is as futile as arguing against Kant that some artworks are *truly* bad, and that certain criticisms are *not at all* arbitrary. These do not disprove the contention that there is a logical problem, they merely make it difficult to identify.

<sup>2</sup> At the same time, it seems that part of Biernacki’s critique is simply that our nets have been cast too narrowly in our haste to run back with the ‘findings’. Thus, his critique of Baumann’s (2007) work on the periodization of film criticism comes from the fact that Baumann constructed a sample from the *New York Times* and concluded that reviews did not treat movies as art until the 1960s, while in fact, such an approach had been common for nearly a half century, simply not in the *Times*. This suggests that a larger sampling frame would have been acceptable, for Biernacki’s (2012b) own conclusions come from his (less formal, but no less real for that) wider frame for collecting his observations (p. 56).

may well be that there is something in the project of science that requires a transposition to the analyst's choice of venue. If so, this cannot be understood as in principle unacceptable for a sociological – a social scientific – analysis of culture.

Yet we recognize that transporting our subject matter to a situation in which we dominate, and one that is out of sight of God and Man, allows us to return with counterfeit. Let us assume that we accept that the brute force of coding does not successfully square the hermeneutic circle, and that the interpretation of 'what this is' for each case remains open. Then we understand that any serious debate about claims based on such codings would require the painstaking attempts at replication that Biernacki made. Biernacki reports being unable to replicate the codings of the sociologists he critiques. But we in no way take this as evidence that there was anything imperfect about their original coding. How could we, *unless we were to replicate the replication* (cf. Steinmetz, 2013, p. 458)? (We noted that the fearsome costs in time of such an act would in themselves be sufficient to explain the lack of engagement with Biernacki's work.) The very fact that we have absolutely no intention of determining whether Griswold's or Biernacki's coding of this or that review is consonant with our own reading is itself the crucial evidence we need to support Biernacki's critique of coding: no science can build objectivity where replication is so implausible. And yet it was a quest for such objectivity that drove the coding movement in the first place.

### **Coding and the halo problem**

Despite this impulse toward impersonal objectivity, coders have long understood the fundamental puzzle of the hermeneutic circle in the practical form of the 'halo' problem. Someone attempting to code *portions* of an interview (whether a clinical psychopathological one or an opinion survey) would come to a decision – often not fully conscious – as to the 'type' of person being coded, which would then bleed into the decisions as to how to code each particularity. When the same particularities were coded independently, by, for example, researchers given disconnected portions of different interviews, the results would be markedly different (see, for example, Hyman and Sheatsley, 1954). (Biernacki (2012a, p. 130) gives clear attention to this problem as endemic to coding.) In many cases, the excessive consistency of halo-based coding quickly approaches stereotype. For example, when a man says it, it is smart; when a woman says it, it is not. In both science and everyday life, when actors want to code honestly, they tend to fight against such halo effects (for example, they institute blind evaluations). When they do not, because blind evaluations will let in too many of the 'wrong sort' of people (for example, too many Asians win violin competitions; inexpensive wines are judged superior to famous and expensive ones), actors will make sure that we do not prevent ourselves from judging the parts in the light of the whole.

Yet we cannot necessarily simply declare war against the halo effect, as to do so would imply refusing to accept the necessity of the hermeneutic circle – attentiveness to context when interpreting any part. The puzzle for us is that this becomes, in potential, radically destabilizing. Borges wonderfully tells the story of Pierre Menard, who attempts to rewrite *Don Quixote* – and succeeds! – but *as a twentieth century Frenchman*. Borges (1964, p. 43) takes a passage from each to compare:

It is a revelation to compare Menard's *Don Quixote* with Cervantes'. The latter, for example, wrote (part one, chapter nine):

... truth, whose mother is history, rival of time, depository of deeds, witness of the past, exemplar and adviser to the present, and the future's counselor.

Written in the seventeenth century, written by the 'lay genius' Cervantes, this enumeration is a mere rhetorical praise of history. Menard, on the other hand, writes:

... truth, whose mother is history, rival of time, depository of deeds, witness of the past, exemplar and adviser to the present, and the future's counselor.

History, the *mother* of truth: the idea is astounding. Menard, a contemporary of William James, does not define history as an inquiry into reality but as its origin. Historical truth, for him, is not what has happened; it is what we judge to have happened. The final phrases – *exemplar and adviser to the present, and the future's counselor* – are brazenly pragmatic.

As always, with his appreciation for the absurd, Borges teaches us something of the dangers of the sublime. Is the best translation of 'yes' sometimes 'no'? Why not, if the most important thing about this 'yes' was that, in the original context, it was a denial of conventional beliefs, and in the target context, beliefs have reversed?

Recognizing the instability that is logically inherent in the hermeneutic project does not imply that interpreters are unable to distinguish the difficult and ambiguous from the horrendous or absurd. However, it does point us to where we need to look for logical problems in interpretation. And doing so will justify our decision not to side with Biernacki (2009a), and refuse to sully 'distinguished humanists' with our own 'crudities' (p. 134).

Consider the case of how to interpret a single writing by the literary critic Paul De Man (we rely on Barish, 2014: 198f). In 1940s Europe, the young De Man, ever the affable opportunist, took charge of a periodical under pro-Nazi control. In it, he penned a stirring defense of modern literature against the accusations of Jewish contamination, arguing that Jewish influence was hardly as great as it might appear, since the Jews often exaggerated their own influence. In fact, he found it 'comforting' that European intellectuals had managed to 'safeguard themselves'

from the 'Jewish spirit'. He even suggested that 'the creation of a Jewish colony isolated from Europe would not entail, for the literary life of the West, deplorable consequences. The latter would lose, in all, a few personalities of mediocre value and would continue, as in the past, to develop according to its great evolutive laws'.

But that was when De Man thought the Germans would win; when it became apparent they would not, he immigrated to the United States, lied about his past and became friends with Jacques Derrida and other literary theorists. Eventually, De Man's early writings surfaced, and even a Heideggerian like Derrida was not pleased to have one of his friends outed as a Nazi sympathizer (and not only because of Derrida's Jewish ancestry). Derrida felt obliged to write *something* about whether De Man's words meant what they seemed to mean, and so he had to put the part (De Man's early essay) in a larger context.

But which context? That of the De Man the embezzler and liar? That of the enlightened friend of Jews? That of the state of Belgium in 1940? Depending on what one picked, it seems that one could read the words rather differently. Certainly, as Derrida (1988, p. 624) argued, if De Man condemns 'vulgar anti-Semitism' (as he did in his piece on literature), it *could be* that he implicitly favors a sophisticated anti-Semitism. But it also could be that he meant to condemn anti-Semitism *tout court*. One must, Derrida insisted, read these words in the context of all De Man's *other* writings (and re-read them, again and again). Still, even with the contextualization, Derrida knew that he was unlikely to be able to convince everyone. How to reach closure?

It is significant that Derrida (1988, p. 591) consistently contrasted the sophistication and complexity of his and De Man's analytic approach to the vulgarity and simplification of De Man's critics, who might be so 'totalitarian' to say something as straightforward as 'De Man is a butthead'. It was De Man's 'rejection of the commonplace' that, Derrida believed, had elevated him over his critics; similarly, Derrida's strategy to support his friend no longer able to mount his own defense was in effect to shame these opponents away from the 'simplism' of recognizing the obvious.

Now it is important to emphasize that Derrida's approach to interpretation is hardly typical for the humanities; further, many humanists would never even raise the question of 'what De Man meant', as it contains too many implausible assumptions about the relation between a text and an author. And thus it is not our intention to tar all 'distinguished humanists' with the brush of Nazism or paradox; our point is simpler and more serious. This sort of interpretation – when we must choose *some* parts in order to make sense of *this* part – is an 'unstable' epistemic strategy, precisely because of the hermeneutic circle, in that the only way one interpretation can be *demonstrated* to be superior to another (another who is not convinced, that is) is through exogenous authority (the person with tenure wins, say). This does not mean that this is *generally* the way in which interpretive disputes are settled, nor even that there cannot be interpretive communities that exist and produce knowledge without *ever* relying on such



authority (for example, they may all be convinced by the same things). Further, we certainly do not deny that such interpretive work can be carried out with rigor, and even with recognizable rigor (such that two efforts may be compared in these terms). To make this claim would be as absurd as taking Kant's third critique to imply that there are no thriving communities of esthetic practice or criticism based on shared understandings.

But this *is* to say that a need for resolution via authority lurks outside the circle of light cast by the fire of successful interpretation, and we should therefore be sympathetic to the search for another way of adjudicating contrasting claims as to the verisimilitude of arguments having to do with cultural artifacts. The reason we do not wish to return interpretation to distinguished humanists, then, is that there are times when the only way they seem to be able to resolve a disagreement is a comparison of which humanist is the more distinguished.

Thus, one problem with humanistic interpretation is that those who are best at it can be so good that they can study texts and find meanings that were not even there.<sup>3</sup> It is entirely possible that conventional standards of a humanistic discipline can be used to orient research away from paradox, incompetence or mere error. Yet, we hold with Kant that there is a sort of proof that is unavailable to the humanist to force an interpretation on a disbeliever, and hence a strong temptation for the invocation of status as a proxy for truth. As Foucault ([1966] (1973)) liked to paraphrase Nietzsche's great question: '*Who is speaking?*'<sup>4</sup>

Thus the drive for rigor and objectivity is, we believe, admirable. Further, there are times when some sort of formalization seems necessary, namely when the range of materials to be examined dwarfs what a targeted reader – and in many cases the analyst herself – can plausibly be expected to master (cf. Mohr and Rawlings, 2012). As said above, in such sprawling cases, we can only interpret one part in the light of tendentiously chosen *other* parts. Thus a formal approach does seem called for in at least some of the circumstances that sociologists of culture are likely to confront. While atomizing a whole seems like a bad thing, the halo effect reminds us that there are problems with *not* atomizing as well.

Even more, what Biernacki (2012a, p. 2) identifies as the core of the formal approach to culture – to 'turn meaningful texts into unit facts for the sake of converting these units back into meanings' – is a defensible and promising approach. Finally, it is not one that implies coding. We go on to examine some of the cases in which formal approaches to culture have proven successful, and see

---

<sup>3</sup> That does not mean that such work is devoid of worth: indeed, we might think that the interpretation of Greek statuary made by Herder (2002 [1768–1770]) – one based on a common misunderstanding of Greek art history – is far superior as an esthetic theory to one that would have been developed by a person who accepted that (as art historians now claim) most of these statues were painted like garish mannequins.

<sup>4</sup> Although Foucault seems to take this discussion out of context, it seems to us wholly in Nietzsche's spirit, for Nietzsche – a trained philologist – also noted that 'The men of corruption are witty and slanderous ... they know that whatever is *said well* is believed' (1974 [1882], p. 97).

what we can learn from these, and then go on to suggest future ways of going further along the road of formalization.

## Successes in Formal Textual Analysis

### The Rorschach test

Biernacki at several points suggests that the formal analysis of culture turns an already meaningful work – which would call out for and constrain interpretations based on its own internal relationships – into a hodge podge of lines and dots, which actually offer the analyst *less* constraint, and make for *more* arbitrary interpretations, than the original. Biernacki likens such interpretation to a Rorschach test. This could well be true in particular cases, and indeed, we must acknowledge that this may be an occupational hazard of formal sociologists with a penchant for descriptive data analysis. However, what is most important about this metaphor is less the tenor than the vehicle. Rorschach tests, once widely given to patients to allow for their projective interpretation, were then interpreted by expert clinicians – whose interpretations here turned out, to the best of the ability of any systematic study to determine, just as self-serving and circular as Biernacki would expect sociologists of culture to be (see, for example, Rosenthal and Rosnow, 1969). The reason is that most of our interpretations are what we might call ‘downwardly permeable’<sup>5</sup> in that they can be influenced by our interpretative scheme. We see in each particularity that which would justify our theory of the whole.

The fundamental problems of false positives in pattern recognition are inherent to any interpretive effort. It is true, we believe, that the de-familiarization that may come from the atomization of cultural products and their re-assembly in a new environment without well-established criteria for interpretation can, just as Biernacki says, increase the risk of such false positives. However, the downwardly permeable nature of informal interpretation means that a formal approach that deals with particularities in atomized fashion may reduce another kind of false positive. That is, formal methods may be extremely good at identifying key features of a text when our normal interpretive mode would be to obscure facts contrary to our ‘feel’ of the whole. Thus turning meaning into facts can, at least in some circumstances, work wonders.

### Authorial identification

One well-understood class of examples of successful uses of such techniques is where we are attempting the ascertainment of authorship. The *Federalist Papers*,

---

<sup>5</sup> This is akin to what Churchland (1988) calls ‘cognitively penetrable’ for the case of visual perception.

written by (or so we think) Hamilton, Madison and Jay, posed a fascinating puzzle for intellectual historians, as it seemed to matter a great deal which paper was written by whom – none were signed, and in some cases, there were conflicting historical claims as to authorship.

Mosteller and Wallace (1963) compared the frequency of the words *by*, *from* and *to* in the disputed papers with a sample of the other writings of Hamilton, Madison and Jay, and found that the three used these terms in different distributions, which could be used to predict the authorship.<sup>6</sup> They went on to use less common (and less frequent) words, the results of which dovetailed with their other analyses, and supported the emerging consensus of more traditional historians that almost all of the disputed papers were written by Madison. (This piece also used a Bayesian approach and negative binomial distribution, which were to make their inroads into common sociological practice a full quarter century or more later.) Other formal techniques applied to the *Papers*, from Rokeach *et al's* (1970) analysis of the use of value terms to Bosch and Smith's (1998) more complex mathematical approach, supported the conclusions of Mosteller and Wallace.

Here there has been relatively little dispute, but the identical process happened, with more furor, in art history. In the late nineteenth century, some art historians, following the lead of Giovanni Morelli, began to try to distinguish forgeries (not copies) from authentic works of masters by paying attention to relatively trivial issues of execution, such as the types of ears that the painter tended to use, and not the more important (and more likely to be imitated) issues of overall composition, style and message (Wind, 1964, p. 35, pp. 38–40). By focusing on details, and having a pre-established set of criteria for identification, the decision was relatively insulated from downward penetration.

The lesson, then, is that there are times when formal techniques may be *superior* to less formal interpretations, and we can guess when these are. They are when parties are interested, when the interpretation requires the integration of widely varying types of data, when there is a high noise-to-signal ratio and when theories of interpretation may have guided the very construction of the works being interpreted. Moreover, perhaps most importantly, such techniques are likely to excel when the interpretation in question is relatively simple, and indeed, less a meaning and more a matter of fact. However, we will go on to argue that this approach can be profitably extended to cases of meaning.

One thing that these formal approaches did was to deliberately *impoverish* the works in terms of their content available for interpretation. Going on the principle that our minds are simpler than we often think, we may find that with

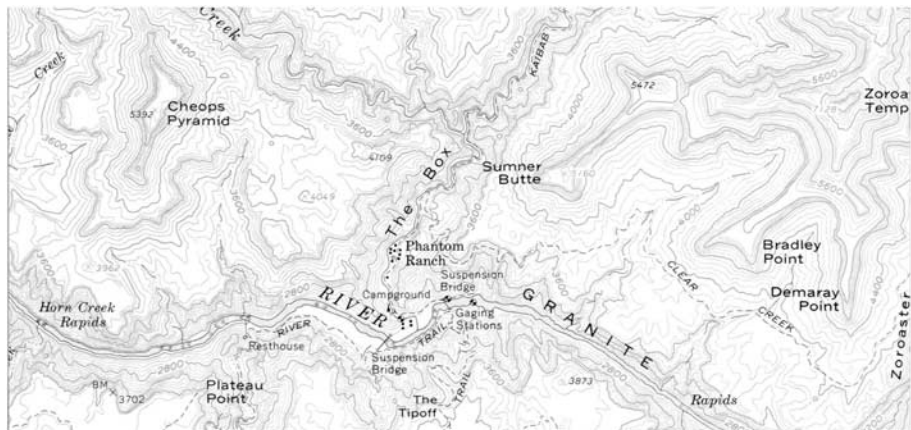
---

<sup>6</sup> Interest in the statistical distribution of letters seems to have begun by Gotthilf Hagen in his 1837 *Grundzuge der Wahrscheinlichkeitsrechnung* (Berlin), which also provided a classic proof of the least squares solution for the normal distribution of error; Lexis (1877, p. 20), like Weber, a member of the *Verein für Sozialpolitik*, used the example of the distribution of letters in Goethe's work to demolish the Queteletian interpretation of statistics that was to be adopted by Durkheim.

too much complex input, we are helpless and unable to know what we are doing and why, let alone determine whether we are right or wrong. It is this deliberate impoverishment of meaning that can make formal techniques successful – but clearly, only successful in a limited sense.

Consider the following chart (Figure 1). This is a conventional contour map of the Grand Canyon. If you have ever been there, you will not fail to notice that this little figure hardly does justice to the Grand Canyon. For one thing, it is much, much, much, much smaller. This changes things: one of the feelings evoked by the Grand Canyon is, in eighteenth century terms, its sublimity: its capacity to dwarf us and make our normal processing powers collapse under its sheer magnitude. One cannot imagine Figure 1 inspiring Ferde Grofe to write the 3"×5" *Contour Map Suite*. Even if we restrict ourselves to more prosaic matters, this map only shows us a top-down (rectangular) projection, while the Grand Canyon can be seen from the North Rim, the South Rim and then from the bottom, not to mention the trails going down the sides.

Given the incredible loss of meaning and information that accompanies the map, why make it all? It is precisely *because* of their impoverishment that maps are useful. First, they can be (and what we call ‘maps’ generally are) substantially smaller than their objects. This is very useful; we can plan a hike down the Grand Canyon in close detail while remaining in our living rooms. In Latourian manner, we reduce the phenomenon to something we can control, and there is no shame in that. Second, the impoverishment is *selective*. As Hutchins (1995) emphasized, maps and charts are usually made to facilitate certain kinds of tasks, and thus suppress certain information to more simply and clearly display others. Figure 1 is a contour map, meaning that it is exceptionally good at giving us a sense of the height (and depth) of the terrain; it is not good for telling us about the vegetation, the soil, how it is used by humans and so on. Third, the impoverishment is



**Figure 1:** Contour map of the grand canyon.  
 Source: Reproduced from a paper original held by the map room, Regenstein Library, Chicago.

*hermeneutically indeterminate*. That is, although an impoverishment may indeed lose information necessary for an ‘interpretation’ – a connection of what we have before us to the meanings we experience as centers of subjectivity – the impoverishment does not force any particular meaningfulness upon us. We can use a map to go many places.

Finally, the map transforms tasks that would otherwise be cognitively difficult into ones that are easy. We may have a very difficult time evaluating whether two lengthy mathematical expressions are equivalent, but a very easy time determining whether two lines are the same length. To the extent that we can replace *quasi*-propositional chains of judgments with visual judgments, we make use of our organic strengths as human beings. How far are two points? Take a piece of string and touch it to the map, move it to the legend. How much higher is one place than another? Trace your finger within a contour from one to as close as you can get to the other, then count the number of lines you traverse before touching the second. The map, then, is an instrument that lets us project complex questions into a reduced space to allow simple cognitive processes to reach adequate understandings that are, within an allowable margin, equivalent to those that could be reached by more cumbersome, and error-prone, processes. This cartographic approach is what we propose for the sociology of culture.

## Cultural Cartographies

### From coding to counting

It is difficult for a formal analysis of culture to do better than the Morellian or Mostellerian approach of classifying place or time of origin. But we believe that new techniques made possible by the digitization of texts allow for an extension of this basic approach. While earlier formal approaches to the sociology of the culture *had* to code in order to go beyond the Mostellerian approach – there was no other way to both construct a plausibly large database and master complexity – technical developments in the computational analysis of text, graphics and sound offer new ways forwards.

The key argument we make is that we can abandon *coding* without abandoning a formal approach. What we substitute is something so simple and vulgar that it might appear incompatible with the technical advance we claim, and it is this: counting. To adapt a cute phrase from Rodney Stark, we are going to argue that when it comes to matters of interpretation, the conclusions of those who count, count.

Recall that the core problems that Biernacki uncovered had to do with coding, and the imposition of interpretations that were (necessarily) tendentious and out of sight of the reader. That is not true for counting, or at least,

not for some sorts of counting, for there are things that we can count that do not require an interpretive act. That is not to say that there are not ambiguities. Imagine that we are sent out to count the squirrels in a park. There may be times when we are not sure whether what we saw was indeed a squirrel. Our number will be imprecise ... but this ambiguity is different from that which arises from the hermeneutic circle, the way it would if we went out in the park to count instances of 'sexism', say. In the latter, *each* time we have a particularity, our decision as to in which class to place it is one that may plausibly be contested, and this is because we may not agree as to what, even in the abstract, sexism 'is', while we have no such disagreement about squirrels.

What are the things we can count? Like Mosteller, we can count words. Now with words (as Mosteller and others noted) there are already more puzzles than we might want. For example, the word 'bar' can be a verb or a noun, and as a noun, it has a number of very different meanings (hence a computer trying to parse the sentence 'we will bar from the bar lawyers found at the bar' might just catch on fire). But note that, unlike with coding, the ambiguity does not affect any decision in the counting process itself ... only our *post*-counting interpretation. Thus, if we find that one text uses the word 'bar' frequently, an English speaker will wonder, which 'bar' was intended? The context will still be required to make such a determination.

We have, then, not yet conducted our act of interpretation when we have made our counts. The interpretation is put off; all we have done is impoverished the cultural artifact in order, as Biernacki has said, to turn it into 'data'. Why? One may argue that we *must* turn the Grand Canyon into an impoverished representation, since no one can contain it in her or his living room, but we do not need to do this with, say, a serious philosophical text. Plenty of us have read these works, multiple times.

And there is no reason why such reading is wrong-headed, but that does not make it sociology. We impoverish to make the artifact 'small enough' so that (again, as Latour, 1986 has emphasized), the writer *and* reader can (metaphorically speaking) both lean over and examine it at the same time, pointing to key features – we can *show*, and not simply *tell*. We do this not with a 'full-size' *part* (as with the selection of an excerpt, a tendentious choice), but with a 'reduced-size' *whole*.

And when we do this, our impoverished representation does not definitively establish any *one* interpretation, but it can shift the burden of proof dramatically. Consider the simplest form of such a reduction. One person argues that Habermas's *Theory of Communicative Action*, often understood to be attempting to deal with the puzzles of agency in system, or that of rationality and ethics, was actually a struggle to refound neo-Kantian philosophy but beginning with Schopenhauer. 'But he does not mention Schopenhauer even once!' we protest. That is a simple, and extreme, version of the counting method in practice.

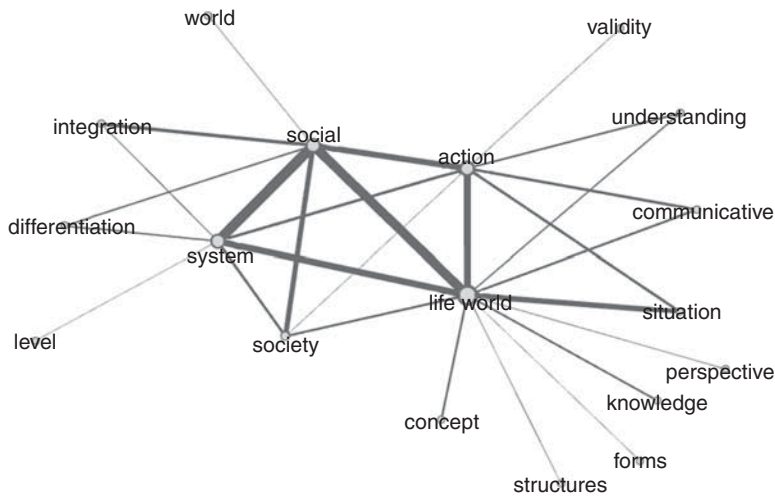


Figure 2: Chapter 6: 'Intermediate Reflections: System and Lifeworld'.

SUM [WORD = 'SCHOPENHAUER'] = 0. Our interlocutor now has a lot more work to do to make this interpretation plausible.<sup>7</sup>

Such summary counts are of course the simplest approach we can take. Things get exponentially more complex when we consider the *adjacency* of words. That is, we count not the number of times any word is used, but the number of times two words occur in close succession (such as in the same sentence or paragraph). Now words can be paired for all sorts of reasons. If, say, we were to find Habermas pairing 'Weber' and 'rationality' that might indeed indicate that Habermas is, if not taking his theory of rationality from Weber, still understanding Weber as the chief interlocutor for a discussion of rationality. However, if we find 'systems' and 'theory' next to one another, this may not mean (or so runs our interpretation) that Habermas sees any theory of society as turning on an analysis of systems; rather, 'systems theory' is a common, two-word (in English), expression with a unitary meaning. And further, if we find 'system' and 'lifeworld' frequently used together, it may not mean that Habermas is saying that the system is the lifeworld or a type of lifeworld – quite the contrary, they are paired as an opposition.

Interpretation, then, is required of the resulting impoverished cartography. So why do it? Consider Figure 2, which is a network composed of the most common word adjacencies in Chapter 6 of Habermas's *Theory of Communicative Action*. (The technical details on the construction of this and other figures are found in the Appendix.) Although this exercise does not shed any interpretive light on this

<sup>7</sup> We note that something very similar recently occurred in debates over the interpretation of Durkheim's work, in Orbach's (1998) critique of Meštrović (1988).

chapter that a competent reader of Habermas would not already have, we note that it does seem to trace the outlines of important connections made by Habermas in the course of his argument.

In Chapter 6, Habermas argues that there are two fundamental ways of conceptualizing the social order: as a *social system* and as a *lifeworld*. While the onset of modernity (and the move away from traditional society) has seen the decoupling of these two conceptions, late modernity has witnessed their reunification, in which the system's instrumental rationality imperatives 'make their way into the lifeworld from the outside – like colonial masters coming into a *tribal society* – and force a process of assimilation upon it' (Habermas, 1987, p. 355). The unification of the *system* and *lifeworld* is represented in the *social-system-society-lifeworld-action* pentagram clique – a merging of the two concepts that is pulled in two different directions by disparate (and thus graphically disconnected) elaborations of the *social system* and *lifeworld*.

A lot of work has gone into producing a formal representation that has neither novelty nor depth; it does seem accurate, on its own impoverished terms, but we would still wonder, why bother? To answer this, let us make an (admittedly exaggerated and ideal-typical) distinction between, first, two *means* of interpreting, formal and substantive. Further, make a parallel distinction between two *objects* of interpretation, form and content. Traditional interpretation is a *substantive* interpretation of *content*. A *formal* interpretation of content would be a purely structuralist account, such as Mische's (2009) idea that what an ideology in a social movement field 'means' is precisely which other groups hold it (see Breiger, 1974). A *substantive* interpretation of *form* happens when we, for example, look at two graphs of belief networks and pronounce one looking 'busier', or when Martin (2002) interprets the negative entropy beyond the marginals of a multiway belief table as Durkheimian 'constraint'. A *formal* interpretation of form is also possible – and this is when we, for example, demonstrate that two seemingly different distributions of public opinion, say, are isomorphic (for example, both having a center-periphery structure; other examples are discussed by Mohr and Rawlings, 2012, p. 86). In *all* cases, including the formal interpretation of form, we still must judge the *meaning* of our results.

Our formal cartography, then, not only simplifies the task of interpretation, and is therefore of use when we would prefer simpler to richer interpretations, if only for reasons of tractability. It also allows for formal interpretations of form, and hence has the equivocal objectivity of such purely formal approaches – it is robustly indifferent to the prejudices of particular investigators ... unless these have been, wittingly or unwittingly, smuggled into the earliest definitions. If such smuggling has not occurred (though of course ontological assumptions are still weighty), such a technique allows straightforward answers to many specifically *sociological* questions that could be extremely cumbersome for a traditional analyst. For example, let us say that we are arguing over whether Habermas or his intellectual heir, Axel Honneth, is truer to the fundamental tenets of critical



theory as proposed by their common predecessors Max Horkheimer and Theodor Adorno. In traditional hermeneutic disputation, this is the sort of question that is very difficult to bring to closure, much to the relief of those of us who have friends or family employed in such endeavors. But sociologists interested in culture – perhaps especially those who adopt what Alexander (2003) sees as the sociology ‘of’ culture, attempting to ‘explain’ its production and dissemination – may be unwilling to wait for the second coming to resolve interpretive ambiguities.

So what if we were to make similar maps, not just for this one chapter, but for the major works of Horkheimer and Adorno, of Habermas, and of Honneth? We first examine each map to make sure that it seems to have validity, in the sense of not suggesting an interpretation that goes against our general understanding of each text.<sup>8</sup> Assuming we find no problems in the individual maps, we then overlay them – and now count up the degree to which, say, the graph of Habermas shares nodes and edges with that of Horkheimer and Adorno, and compare that degree with the degree to which *Honneth’s* graph shares nodes and edges with the earlier critical theorists. If the number indicating overlap between Habermas and Horkheimer is greater than that for the Honneth–Horkheimer overlap, we would have *prima facie* evidence of a declining commitment to these core values.

Or that is how most of us would interpret these numbers. We might not: there could be all sorts of reasons why we would refuse to put our faith in the results of such a procedure. Note that while it may *appear* that we have been able to carry out our systematic exploration without needing to interpret the content of the texts at all, this is not quite so. We have simply made a class of assumptions of interpretive equivalence which we have not actually checked – assuming (for example) that the meaning of the relation between two words in one graph is equivalent to the meaning of the relation between the same words in a *different* graph. This assumption may be incorrect, and it may well be that a critic who suspects this will need to return to substantive interpretation of the content to assess this assumption. The method is not foolproof, but it also lacks identifiable biases in favor of the analyst’s preconceptions. Although the matter is not closed, the burden of proof should have shifted – someone who wants to argue the opposite would need to determine why, if our interpretation is wrong, we see the

---

<sup>8</sup> We note that if, after performing our formal analysis, we decide that it gives us a perverse interpretation, and hence (we conclude) must contain problematic assumption, and we therefore revise our formal procedures, then we *do* have a back-stage imposition of interpretation, precisely the situation we are attempting to avoid, as the researcher’s preconceptions are built into the procedure. This means that formal procedures, including the ones we use here, that are adapted too many times, or that have too many ‘tunable parameters’ other than obvious things like width of view, or focus, do not achieve the results we want. However, we have been pleased to find that in our case, our procedures do not rely on such weeding out of the seemingly perverse contradictions to our preconceptions. They worked fine, first crack off the bat.

pattern in the data that we do (Latour, 1987). It is not enough to point to the limitations and impoverishment in the procedures – all sociological data are incredibly impoverished, and they are subject to all sorts of differential influences, but we have learned that in enough cases, when we throw it all together, the important principles of the world will shine through.

Now here is the funny thing – Lee actually did the analysis laid out above, and found that it was quite the opposite – it is Honneth, not Habermas, whose concept map (of his early work) agrees with the approach of Horkheimer and Adorno. Figure 3 summarizes her findings, here focusing on the subgraph of each

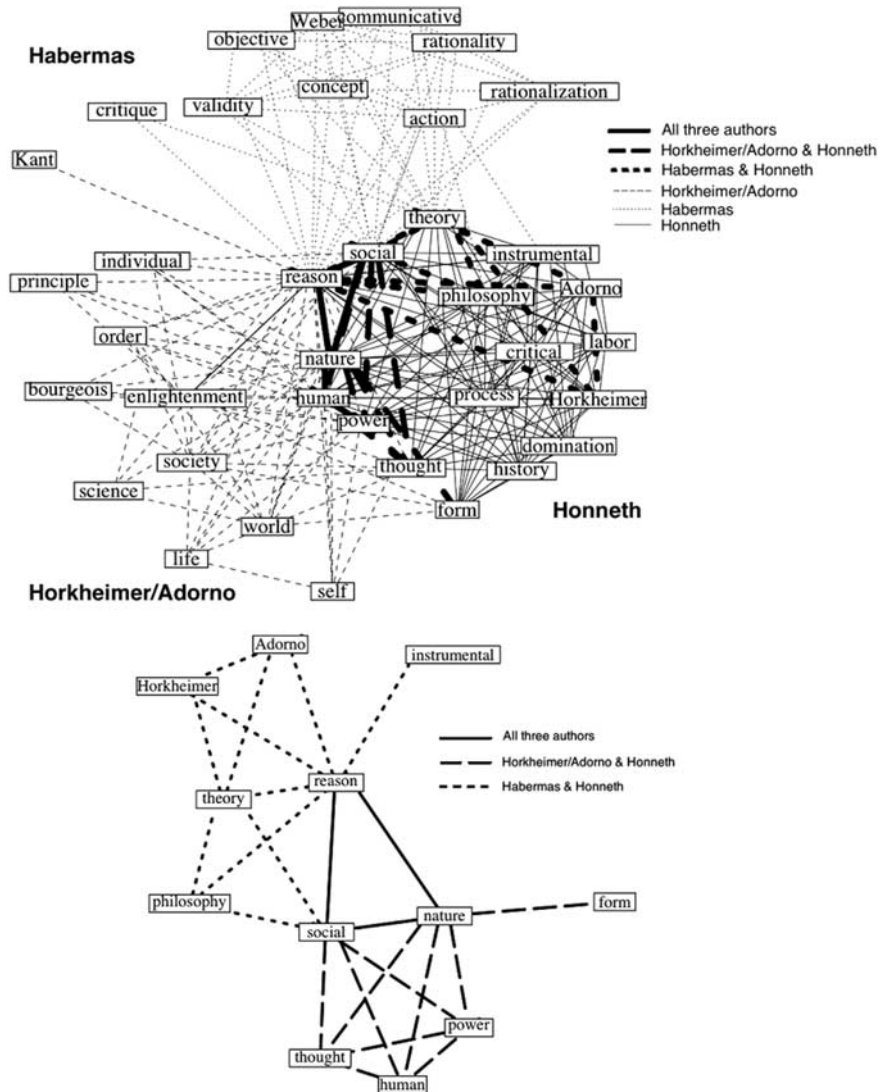


Figure 3: Overlaps of concept maps of three generations.

that is tied to the concept 'reason'. Below the main graph is a recasting that only focuses on the edges that are part of an intersection between the three main subgraphs. While there is ample overlap between Honneth and Habermas's as well as Honneth and Horkheimer/Adorno's thoughts on the key Critical Theory concept of 'reason', there is scarcely any overlap between Habermas and Horkheimer/Adorno's ideas. They share only three edges, all of which are present in Honneth's map as well. This fits Lee's own reading of Honneth's work as a return to some of the first generational principles, and even more, Honneth (in an interview with Lee) confirmed that this had been his deliberate goal. Of course, a subjective interpretation may not be shared by a reader, and an author's account of his own motivations is always suspect, but it is extremely gratifying that the formal method, again, agrees with other interpretations in a matter in which one can imagine a much wider range of opinion.

And we can go even further. We can compare the way in which two authors use a third. Erich Fromm was once a member of the Frankfurt school, but was basically expelled because of his ideational disagreements with the others. Most important, or so say intellectual historians (see Burton, 1991; Wiggershaus, 1995; Kamau, 2012), Fromm's interpretation of Freud differed from that of the other theorists, especially Marcuse, who increasingly drew upon psychoanalysis. Did Fromm use Freud in one way, and the other critical theorists another?

By looking at the 100 terms most likely to be used in a paragraph with the word 'Freud', we actually find that Adorno's deepest engagement with Freud, his 1961 *Notes to Literature 2*, has only a minor (8 per cent) overlap with Marcuse's pivotal engagement with Freud (his 1955 *Eros and Civilization*), while Marcuse's work has a substantial (30 per cent) overlap with Fromm's 1941 *Escape from Freedom* – the work where Fromm's use of Freud purportedly prompted his expulsion from the school. (Readers will find a more complete discussion of the procedures used here in the Appendix.) Even more, Fromm's successive works maintain this overlap with Marcuse in key terms associated with Freud; it is identical in Fromm's 1955 *The Sane Society* and his 1973 *Anatomy of Human Destructiveness*. And if we look at which of these common terms are *tied* to one another, we find that in this last work, fully 30 per cent of the *edges* connecting these shared concepts are present in both Marcuse's graph and Fromm's – a higher overlap than those seen between Fromm's own works.

In other words, Fromm continued to circle around the same basic concept structure that formed the core of Frankfurt's orthodox view of Freud, as opposed to using a wholly 'different' Freud. And it was, it seems, this *overlap*, and not difference, that caused Fromm's expulsion (similar dynamics are found in the history of other one-time members). Thus the key thing in these maps is *not* the pictorial representation, which is only of use insofar as it allows us to check that we have not made an obvious mistake (such as treating a word and its plural as two different words, or, contrarily, collapsing homonyms). Rather, it is that the graphs are subject to manipulation that can lead to quantitative results that have

still evaded interpretive closure (and it is only the numbers that must now be interpreted).

Finally, it is worth emphasizing that the simple maps used here have no structure – every word is treated as equivalent, with the exception of the clutter of ‘stop words’ like *a* and *but*. But algorithms to find grammatical position are encouraging (see, recently, Mohr *et al*, 2013) if still naïve, suggesting that we have only scratched the surface of the sorts of reduced cartographies made possible by formal methods.

It is worth re-emphasizing that it is by no means obvious that, even in the sociology of culture, such a formal approach always brings with it sufficient advantages to justify its use. If a sociologist making a claim about late Elizabethan gender relations employs an interpretation of Viola in *Twelfth Night*, it seems hardly necessary to use natural language processing. The text is there, you may make your interpretation, and if the reader disagrees, she may pull her copy down and look at the context of the words cited. But if an author were to try to make claims about texts that the reader does not have on her shelf, or that involve a huge range of texts (such as ‘Elizabethan drama’), it might prove impossible, as a sociologist, to make much headway. It is when interlocutors do not have access to source materials, either because they are unfamiliar with them, or because the size of the corpus dwarfs sociologists’ sense of a plausible time commitment, that we should move toward formalization that can simplify the interpretive task, making it more systematic, because there seems to be no other way for author and reader to be thinking about the same things.

These techniques, as Mohr and Rawlings (2012, p. 73) wonderfully put it, have a necessary ‘brutalization of reality’. We think that it would be absurd to deny that there is a cost to this condensation of reality. These techniques exchange detail for robustness and, frankly, speed. We do not think that they can replace interpretation in humanistic fields, but they may be adequate to answer many sociological questions, and allow for decent research that has the characteristics that serious scholars would demand.

In sum, our argument is that there *are* formal procedures that do not have the problems identified by Biernacki, and that we should substitute cartography for coding. Although we do not here argue that we *should* substitute cartography for traditional interpretation, but rather merely acknowledge that there are reasons why sociologists will often find traditional interpretation incompatible with their training, goals or questions, we also do not deny that the formal techniques we propose may be able to surprise us by suggesting new interpretations even of material that we are able to approach via non-formal methods.

Finally, we recognize that such a cartographic approach is subject – and presumably *will* be subject – to Biernacki’s (2012b) deep and important criticism of the very idea of dimensionalization. But what sociologists need now are not *perfect* techniques, but *better* ones.

## For the Future

### Recapitulation

We have argued that Evans, Griswold, Bearman and Stovel were great explorers pushing out into the unknown world of formalization. We seek to continue in their footsteps. Yet we also believe that the core arguments of Biernacki were correct. They are, to recapitulate, as follows: cultural analysis *always* involves interpretation. The decision of what to study already builds in an interpretive bias, and involves circularity. And the process of coding violates the canons both of humanist/interpretive and sociological/formal analysis, by forcing an interpretation upon the materials ‘in the dark’, where it is out of the light of common scrutiny, by defining what should be demonstrated, and by (in our terms) combining ideal-typological ways of thinking with nominalist ways of working. Our analysis also implies that if coding is actually possible in an unproblematic form – in the sense that determinate rules may be specified for how each case is treated – then the labor may safely be turned over to computers, and hence *should* be, with readers able to see the computer’s instructions.

But more important, we have argued that there is an approach to formalization that does not involve coding, and indeed, does not necessarily require the construction of a ‘database’ that is any different from what a non-formal interpretive approach would use (for example, focal texts). This approach involves a radical dumbing down of cultural works, and so may at first seem a step backwards. But the reason that hermeneutics remains open ended is because it is comfortable in pastures of verdant complexity in which things can turn into their opposites. We share the belief of many social analysts that non-formal interpretation is *not* inherently up-for-grabs, and we reject the idea that there is no better or worse in such interpretation. Our argument is not that *only* a formal approach can order such a field. However, it is our claim that what *sociology* can offer the study of culture is this formal approach, and that, moreover, this is a specifically *scientific* approach.

Biernacki has skewered the pretensions to science that many formal analyses have. It may sound quaint at best, or the nadir of oafish and clueless vulgarity at worst, to actually claim the mantle of science. But we make this explicit link not to claim the moral high ground, by identifying with a *prescriptive* notion of ‘true science’. Rather, we argue that the approach we lay out here does, in fact, possess key features that have been identified by *descriptive* analyses of science by social scientists – analyses that have often been understood as *undermining* the normative accounts.

We agree with Latour that science involves making phenomena small to the point where they may be mastered: recombined, transferred to paper and made the subject of joint visual attention of (often physically copresent) groups of experts. Were this the only form of mental life that we had, we should all be the

worse for it. But it is a remarkably effective form of weeding out some statements, and these are, we believe, statements that are disproportionately likely to *deserve* weeding out. Put another way, if science is (as Hull, 1988 and others have argued) an evolutionary process, one of the reasons why it is so successful is that most of us leave no descendants. One of the problems with the social sciences is our *lack* of extinctions: a wrong theory, charmingly put, will always find new 23 year olds to recruit (see Nietzsche's point in Note 4). We are confident that Biernacki will agree with us that some ideas are bad ideas. And in the sociology of culture we have many bad ideas that only the violence of science can silence. We do not mince words here, as our aim is to lay out, in clearest form possible, the issues that confront us. As Weber would say, once the situation is clear, it is only 'big children' who refuse to face it and make a choice.<sup>9</sup>

### Contemporary methods

An implication of our approach is that a number of new formal techniques do *not* avoid the problems Biernacki identified. For example, machine learning can be used to 'train' a computer to code documents. The researcher reads through, finds a few examples that illustrate what she is looking for, feeds them into the computer and asks the computer to identify those that are like each example. This is not unlike asking the computer to identify who is a 'hippie' by inserting a picture of Charles Manson, and then treating all those such identified as Mansonites. The key interpretation is imposed by fiat, and is not subject to scrutiny, as in the procedures we advocate. While such methods (see, importantly, Hanna, 2013) can be vital for tasks that involve the consultation of large amounts of primary materials, generally structurally simple ones, such as newspaper articles, and may in fact do so without degradation of quality compared with other plausible methods, they do not solve the specific problems that have been our focus here.

Similarly, methods that 'cluster' observations into clumps and then allow the researcher to *label* the clumps (such as topic modeling), also may (*may* – this is not a necessary result) share the fundamental corruption of nominalistic coding procedures. As such procedures rely fundamentally on counting word adjacencies like the approach we suggest, they escape the self-fulfilling prophecy of imposing a presumed coding scheme on a text corpus. But since clustering takes the further step of unambiguously sorting texts or bits of text into bins (clusters or topics), it replaces the straightforward map of words' relations to one another with an array of categories whose discovery is at once largely dependent on the population of texts included in the corpus and mostly obscure to the reader (and

---

<sup>9</sup> It is not clear to us, we note in passing, whether Biernacki agrees with Weber's perhaps most fundamental argument – that a devotion to science requires abandoning, at least in scientific efforts, a commitment to other values such as the esthetic. 'You serve one God, and offend the others'.

indeed often to the researcher herself). If we are to use such methods, they must be, in the words of DiMaggio *et al* (2013, p. 577), not only *automated* and *inductive*, but '*explicit*, so that the data are available ... for other researchers to reproduce the analyses'. But it is still a challenge to present the data in a form that a reader can validate.

Clustering methods therefore may easily become a form of coding; while in traditional coding, the interpretive scheme is formed first (though usually after some informal reading of the data), and then the partition of the data takes place accordingly (and deductively), here the partition of the data takes place first, and then interpretation afterwards. It is for this reason that Mohr and Bogdanov (2013, p. 560) emphasize that 'well informed interpretive work – hermeneutic work – is still required'. Good topic modelers will present the detailed lists of words associated with each topic, to give the reader some understanding of the logic whereby a *name* is affixed to a cluster. However, to the extent that analysts also rely on their wider reading of (some of) the texts in question, we may have that form of out-of-sight interpretation and *ex cathedra* imposition of interpretations Biernacki criticizes in coding, despite researchers' best intentions.<sup>10</sup> Most important, we must beware of any technique whose results fundamentally depend on what we choose to *call* our categories, and topic modeling is not immune to this problem.

Put another way, such methods, like factor analysis or latent class analysis, may well be revealing, and allow for a reduction of complexity. Further, when analysts provide the detailed information about the terms that load most heavily on the algorithm's typologizing, they make possible just that 'open air' interpretation that we advocate. But such methods share the fundamental duality (in the sense of Breiger, 1974) of most of our methods – we make factors, variables or clusters, based on the cases, and then use these factors, variables or clusters, to 'explain' the cases. The explanation can easily be *too* good, because we are, in effect, paying ourselves with the money we borrowed (from ourselves). Such models may be used to create novel findings about the relation of texts to one another *independent* of the labels attached to the categories – which we see as the virtue in the approach – but it may be difficult to have such methods avoid simply going backwards (we make topics on the basis of the overlap of texts, and then explain the overlap of texts on the basis of the 'topics'). The approach we have outlined here is different, in that our more highly leveraged conclusions do not involve a reparameterization of the duality in the basic data matrix.

---

<sup>10</sup> Indeed, it is because of this openness that careful researchers like McFarland *et al* (2013, p. 613) must rely on exogenous authority to determine what they call 'ground truth', as 'topics were accepted and rejected on the grounds of whether two experts recognized and agreed that the identified word-sets constituted research areas'. There is certainly nothing wrong with the use of accumulated credibility to train new researchers, whether human or non-human; our only point is that this validation in topic models may need to happen backstage.

Thus, although we admit that such clustering efforts will prove necessary for certain substantive problems, our interpretations of these cannot place analytic weight on what the researcher *calls* the resulting categories. When it comes to formal analyses, we might say that bad sociologists code, and good sociologists count. The reason is that the former *disguises* the interpretation and moves it backstage, while the latter *delays* the interpretation, and then presents the reader with the *same* data on which to make an interpretation that the researcher herself uses. Even more, the precise outlines of the impoverishment procedure is explicit and easily communicated to others for their critique. And it is this fundamentally shared and open characteristic that we think is most laudable about the formal approach.

### Humanism

Biernacki clearly not only wants us to return interpretation to the care of humanists, but he considers these humanists to exist in a stratified world away from the ‘crudities’ of the mass of (in some ways) less well-educated and certainly less sophisticated sociologists. We think that this would be a disaster for sociology, as it would be for science in general. While sociologists are used to emphasizing the importance of hierarchy in science, it is worth comparing this form of organization to that which was found in academic environments before the scientific revolution, when dissent was thought crime. (Thus, the motto of the seventeenth century Royal Society: To swear allegiance to the words of no master (Gaukroger, 2006, p. 353).) We do not intend to strip humanist scholars of their positions, but we believe firmly and without apology that scientific approaches – or at least those that find their testing in the open sun of shared debate – have much to offer.

Should sociologists of culture follow Biernacki in attempting to defend ‘humanistic interpretation’ from ‘impertinent’ coding (2014, p. 179), and ‘to segregate quantitatively inspired “scientific” procedure from interpretations of texts and culture’ (2012a, p. 8) – that is, to maintain the division between sacred and profane? We say no, and here, at any rate, side with Durkheim, he who impressed upon sociologists the importance of this division. In his lectures on *Moral Education*, Durkheim (1961 [1902–1903]) considered the argument that some things – morality itself – were too sacred to be subjected to positive knowledge. He rejected these arguments, and we reject them as well.

Biernacki (2012a, p. 58), though continually appealing to Weber as having already solved the most important problems of interpretation,<sup>11</sup> notes that in

---

<sup>11</sup> Biernacki (2014, p. 185) hints that the problem of how to connect evidence to claims of meaning via ideal types had already been solved by Weber, ‘without moving into post-modern verbosity’. We agree to some extent, but would amend it that Weber was content to remain in high-modern verbosity. On the issue of interpretation and its relation to objectivity, Weber says (Biernacki cites the *Objectivity* essay, we that on *Knies*), ‘Diese Objectivierung wird sich, wo es sich um die Ausnutzung



*The Protestant Ethic*, Weber's reading of Franklin was abysmally wrong. Combining a vision of Philadelphia (the largest city in the United States) as some sort of *Hee Haw* set, and a hilarious uncritical acceptance of the mealy mouthed phoney baloney that Franklin printed to sell almanacs, Weber's interpretation of Franklin's ethic – the lynchpin of the argument in *Protestant Ethic* – was an absurdity. In contrast, Weber's interpretation of Luther, though less original, has stood the test of time. It is perhaps significant that the key piece of evidence Weber (1976 [1920–1921], p. 79) used here was a word count (albeit of a simple kind): the first time the word *Beruf* was used to mean 'occupation' was in Luther's translation of the Bible into the vernacular.

Interestingly, it was such close examinations of words used in the Bible, and its translations, that were both the foundation of hermeneutics and, as Biernacki notes, the cause of a debate similar to that unfolding here, as in the eighteenth and nineteenth centuries, traditional hermeneutics became challenged by historical criticism of the Bible. There were three sides: the good – those vested with authority with which to mete out spiritual, if not temporal, punishments to anyone who dared to challenge their interpretations; the bad – the new audacious interpreters who could persuade readers through their brilliance and grace; and the ugly – the historians who paid attention to the way in which cognates of Aramaic words were used in various texts that had been passed down, who poured over the close details of which word was used for God at which point in the text, who compared variants and who counted. We side with those who think that the ugly gave us the most beautiful interpretation of all – that which we believe to be empirically defensible.

## Acknowledgements

We are grateful to three anonymous reviewers and the editors for critical comments that increased the cogency of our contribution.

---

unserer Fähigkeit des »deutenden« *Verstehens* handelt, teilweise, namentlich in der Art und Weise ihrer begrifflichen Bestimmtheit, anders gestalteter Demonstrationsmittel bedienen, als da, wo das Zurückgehen auf »unverstandene« aber eindeutig bestimmte »Formeln« das Ziel sein *soll*, und allein sein *kann*, aber »Objectivierung« ist sie eben auch'. Compressed translation: 'establishing objectivity when it comes to meaningful interpretation is different from that for non-meaningful things, but that doesn't mean it's impossible' (Weber 1922, p. 89). Such problems, wrote Weber, 'can be answered only by a *theory of "interpretation,"* a theory which at this point is barely visible and has hardly been explored at all. For the present and in relation to the foregoing discussion, I shall attempt only to establish the status and possible importance of this problem *as I see it*' (1922: 91f; here we give the translation of 1975 [1903–1906], p. 151). We thus are less sanguine than Biernacki (2012b: 64f) that Weber will supply us with the answers that we need. Further, like many others, we believe ideal-typification to be a completely untenable approach even for rigorous humanistic interpretation (though it may have uses for a science of action), and we look forward to seeing a development here that proves us wrong.

## About the Authors

Monica Lee is a Doctoral Candidate and William Rainey Harper Dissertation Fellow at the University of Chicago. Her research, which examines the history of Frankfurt School philosophy via computational methods of text analysis, was performed as a visiting scholar at the *Frankfurt Institut für Sozialforschung* and was supported by the DAAD, the Leo Baeck Institute and the *Studienstiftung des deutschen Volkes*.

John Levi Martin is a Professor of Sociology at the University of Chicago. He is the author of *Social Structures*, *The Explanation of Social Action* and *Thinking Through Theory*, in addition to articles on the formal analysis of subjective structures.

## References

- Alexander, J.C. (2003) *The Meanings of Social Life*. New York: Oxford.
- Barish, E. (2014) *The Double Life of Paul De Man*. New York: Norton.
- Baumann, S. (2007) *Hollywood Highbrow: From Entertainment to Art*. Princeton, NJ: Princeton University Press.
- Biernacki, R. (2009a) After quantitative cultural sociology: Interpretive science as a calling. In: I. Reed and J.C. Alexander (eds.) *Meaning and Method*. Boulder, CO: Paradigm Publishers, pp. 119–208.
- Biernacki, R. (2009b) The banality of misrepresentation. In: I. Reed and J.C. Alexander (eds.) *Meaning and Method*. Boulder, CO: Paradigm Publishers, pp. 253–262.
- Biernacki, R. (2012a) *Reinventing Evidence in Social Inquiry*. New York: Palgrave Macmillan.
- Biernacki, R. (2012b) Rationalization processes inside cultural sociology. In: J. Alexander, R. Jacobs and P. Smith (eds.) *The Oxford Handbook of Cultural Sociology*. New York: Oxford University Press, pp. 46–69.
- Biernacki, R. (2014) Humanist interpretation versus coding text samples. *Qualitative Sociology* 37(2): 173–188.
- Borges, J.L. (1964) Pierre Menard, author of the Quixote. In: D.A. Yates and J.E. Irby (eds.) *Labyrinths*. Translated by James E. Irby. New York: New Directions, pp. 36–44.
- Bosch, R.A. and Smith, J.A. (1998) Separating hyperplanes and the authorship of the disputed federalist papers. *American Mathematical Monthly* 105(7): 601–608.
- Breiger, R.L. (1974) The duality of persons and groups. *Social Forces* 53(2): 181–190.
- Burton, D. (1991) *The Legacy of Erich Fromm*. Cambridge, MA: Harvard University Press.
- Churchland, P.M. (1988) Perceptual plasticity and theoretical neutrality: A reply to Jerry Fodor. *Philosophy of Science* 55(2): 167–187.
- Derrida, J. (1988) Like the sound of the sea deep within a shell: Paul de Man's war. Translated by Peggy Kamuf. *Critical Inquiry* 14(3): 590–652.

- DiMaggio, P., Nag, M. and Blei, D. (2013) Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics* 41(6): 570–606.
- Durkheim, E. (1961 [1902–1903]) *Moral Education*. Translated by Everett K. Wilson and Herman Schnurer. New York: The Free Press.
- Evans, J. (2009a) Two worlds in cultural sociology. In: I. Reed and J.C. Alexander (eds.) *Meaning and Method*. Boulder, CO: Paradigm Publishers, pp. 209–252.
- Evans, J. (2009b) Imperviousness to disconfirming data. In: I. Reed and J.C. Alexander (eds.) *Meaning and Method*. Boulder, CO: Paradigm Publishers, pp. 263–277.
- Foucault, M. ([1966] (1973)) *The Order of Things*. New York: Vintage Books.
- Fromm, E. (1941) *Escape From Freedom*. New York: Farrar & Rinehart.
- Fromm, E. (1955) *The Sane Society*. New York: Rinehart and Company.
- Fromm, E. (1973) *The Anatomy of Human Destructiveness*. New York: Holt, Rinehart & Winston.
- Gadamer, H.G. (1975 [(1965)] *Truth and Method*. Translated by Garrett Barden and J. Cumming from the 2nd edn. New York: The Seabury Press.
- Gaukroger, S. (2006) *The Emergence of a Scientific Culture: Science and the Shaping of Modernity, 1210–1685*. Oxford: Oxford University Press.
- Habermas, J. (1987) *The Theory of Communication Action*. Volume 2, translated by Thomas McCarthy. Boston: Beacon Press.
- Hanna, A. (2013) Computer-aided content analysis of digitally enabled movements. *Mobilization* 18(4): 367–388.
- Herder, J.G. (2002 [1768–1770]) *Sculpture*. Translated by Jason Gaiger. Chicago, IL: University of Chicago Press.
- Hull, D.L. (1988) *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago, IL: University of Chicago Press.
- Hutchins, E. (1995) *Cognition in the Wild*. Cambridge, MA: The MIT Press.
- Hyman, H.H. and Sheatsley, P.B. (1954) ‘The authoritarian personality’ – A methodological critique. In: R. Christie and M. Jahoda (eds.) *Studies in the Scope and Method of The Authoritarian Personality*. Glencoe, IL: The Free Press, pp. 51–123.
- Kamau, C. (2012) On Erich Fromm: Why he left the Frankfurt School. In: D. Berry (ed.) *Revisiting the Frankfurt School: Essays on Culture, Media and Theory*. Surrey, Canada: Ashgate, pp. 189–196.
- Kant, I. (1987 [1790]) *Critique of Judgment*. Translated by Werner S. Pluhar. Indianapolis, IN: Hackett Publishing Company.
- Kuyvenhoven, C. (2013) Book review: Richard Biernacki, reinventing evidence in social inquiry. *Canadian Journal of Sociology* 38(2): 292–294.
- Latour, B. (1986) Visualization and cognition: Thinking with eyes and hands. *Knowledge and Society* 6: 1–40.
- Latour, B. (1987) *Science in Action*. Cambridge, MA: Harvard University Press.
- Latour, B. (1999) *Pandora’s Hope: Essays on the Reality of Science Studies*. Cambridge, MA: Harvard University Press.
- Lexis, W.H.R.A. (1877) *Zur Theorie der Massenerscheinungen in der menschlichen Gesellschaft*. Freiburg, Germany: Wagner.
- Martin, J.L. (2002) Power, authority, and the constraint of belief systems. *American Journal of Sociology* 107(4): 861–904.

- Martin, J.L. (2014) *Thinking Through Theory*. New York: Norton.
- McFarland, D.A., Ramage, D., Chuang, J., Heer, J., Manning, C.D. and Jurafsky, D. (2013) Differentiating language usage through topic models. *Poetics* 41(6): 607–625.
- Meštrović, S. (1988) The social world as will and idea: Schopenhauer's influence upon Durkheim's thought. *Sociological Review* 39(4): 674–705.
- Mische, A. (2009) *Partisan Publics: Communication and Contention across Brazilian Youth Activist Networks*. Princeton, NJ: Princeton University Press.
- Mohr, J.W. and Bogdanov, P. (2013) Topic models: What they are and why they matter. *Poetics* 41(6): 545–569.
- Mohr, J.W. and Rawlings, C. (2012) Four ways to measure culture: Social science, hermeneutics, and the cultural turn. In: J. Alexander, R. Jacobs and P. Smith (eds.) *The Oxford Handbook of Cultural Sociology*. New York: Oxford University Press, pp. 70–113.
- Mohr, J.W., Wagner-Pacific, R., Breiger, R.L. and Bogdanov, P. (2013) Graphing the grammar of motives in natural security strategies: Cultural interpretation, automated text analysis and the drama of global politics. *Poetics* 41(6): 670–700.
- Mosteller, F. and Wallace, D.L. (1963) Inference in an authorship problem. *Journal of the American Statistical Association* 58(302): 275–309.
- Nietzsche, F. (1974 [1882]) *The Gay Science: With a Prelude in Rhymes and an Appendix of Songs*. Translated by Walter Kaufmann. New York: Vintage.
- Orbach, H.L. (1998) The supposed influence of Schopenhauer on Durkheim. Anatomy of a modern myth that exemplifies Merton's establishing the phenomenon. *Soziale Welt* 49(1): 71–90.
- Riley, D. (2014) Back to Weber!. *Contemporary Sociology* 43(5): 627–629.
- Rokeach, M., Homant, R. and Penner, L. (1970) A value analysis of the disputed federalist papers. *Journal of Personality and Social Psychology* 16(2): 245–250.
- Rosenthal, R. and Rosnow, R.L. (eds.) (1969) *Artifact in Behavioral Research*. New York: Academic Press.
- Steinmetz, G. (2013) Cultural coding's condemner condemned. *European Journal of Sociology* 54(3): 453–460.
- Weber, M. (1922) *Gesammelte Aufsätze zur Wissenschaftslehre*. Tübingen, Germany: J C B. Mohr.
- Weber, M. (1975 [1903–1906]) *Roscher and Knies: The Logical Problems of Historical Economics*. Translated by Guy Oakes. New York: The Free Press.
- Weber, M. (1976 [1920–1921]) *The Protestant Ethic and the Spirit of Capitalism*. Translated by Talcott Parsons. New York: Charles Scribner's Sons.
- Whewell, W. (1971 [1860]) *On the Philosophy of Discovery*. New York: Burt Franklin, Orig. Publishers, Reprinted from original by Lenox Hill Publishers, New York.
- Wiggershaus, R. (1995) *The Frankfurt School: its History, Theories, and Political Significance*. Cambridge, MA: MIT Press.
- Wind, E. (1964) *Art and Anarchy*. New York: Knopf.

## Appendix

Here we lay out the technical details of the procedures used to make the maps and related analyses used in the text.

### Choice and processing of texts

Some of the texts analyzed here were originally written in German, but we here use English translations; this is a choice made for reasons of precision, not of convenience. A semantic relational analysis based on copresence in semantic neighborhoods is most reliable when each concept's symbolic representation is consistent throughout a text; that is, a concept is expressed in as few forms as possible – Concept A is represented only by word A, not by A as well as *Ar*, *An* and *As*. Owing to its conjugation and declension rules, as well as the tendency of writers to create compounds, concepts are relatively inconsistently symbolized in German. We are therefore likely to achieve more reliable results by analyzing our texts in English, given that the texts here were translated with great care. Further, the pivotal concepts here are ones of recognized philosophical import, and thus, like other 'technical terms', usually kept consistent throughout the translated text.

The texts were digitized by being scanned and subjected to optical character recognition software; the quality of the text scans was such that there were very few errors in terms of conversion to text. A second program then dropped what are called 'stopwords' – articles, many prepositions and so on – to allow a semantic analysis to focus on the theoretically important terms.

### Nodes and edges

Following the procedures developed by Lee, we do not collapse different forms of words because investigation of the texts suggests that in general these have different usages (thus a philosophical discussion of 'social' is different from one of 'socialization'). (The exception is that singulars and plurals are combined unless there is an *a priori* reason to consider them different.) The most commonly used words then can be chosen for graphical display.

Second, we may establish a symmetric tie between two terms if they are used textually near one another. Exhaustive investigation by Lee suggested that the paragraph unit was most likely to produce results that corresponded to an informal reading. This is true even for texts that were originally written in German with far fewer paragraph breaks; English translators often add paragraph breaks not present in the original to separate ideas, and thus increase the validity of the results, though this does imply that the translator's interpretations have some degree of influence over the establishment of a conceptual structure. We then have data that are suitable for being treated in graph theoretic form,

where a graph is a set of ‘nodes’ (here words) and ‘edges’ (here relations between words).

We note that here we do not apply a conventional inverse frequency weighting to the terms, such as is often used when we are interested in the *association* between any two terms chosen at random (or all such pairs of terms). The reason for this choice is that such an approach, which conditions away the marginals, throws out the most important information for a cartographic analysis, namely, the information as to which terms are most frequent. Such weighting, we have found, fragments graphs and elevates trivial terms to central positions. The unweighted versions better reproduce the core arguments, at the cost of the edges lacking an interpretation as an associational parameter.

### Creation of concept maps

With these counts, we can create what may be called ‘concept maps’, networks of ties between simple ideas (here words). What we call a *semantic relation* is defined as the copresence of words in paragraph. As said in the text, the precise *nature* of this semantic relation is *not* established by the procedure but requires interpretation.<sup>12</sup> We can weight these edges by their frequency; for example, Table A1 (infra) has a list of semantic edges summarizing Chapter 6 of Habermas’s *Theory of Communicative Action*. Edge weights range from 130 to 384, meaning that each pair of words is copresent between 130 and 384 times over the course of the chapter.

Once such a concept map is created, we are free to focus on subsets of this (subgraphs) for particular theoretical issues. Figure 3 in the article uses three texts (Horkheimer and Adorno’s *Dialectic of Enlightenment*, Habermas’s *Theory of Communicative Action* and Honneth’s *Critique of Power*), turned into concept maps as discussed above. From each book’s concept map, the subgraph containing the 16 words most strongly tied to the term ‘reason’ was created. Aforementioned Figure 3 then shows these three graphs overlaid. If an edge is present in more than one work (technically, such an edge is in the ‘intersection’ of the two subgraphs), it is drawn in a distinctive fashion. Fourteen are shared between Honneth and Habermas, and 21 are shared among Honneth and Horkheimer/Adorno. But only three edges are shared between Habermas and Horkheimer/Adorno, and these are also shared by Honneth as well. Table A2 (infra) shows the semantic ties shared among Horkheimer/Adorno, Habermas and Honneth.

In the case above, the focus on the concept *reason* was made because of specific arguments in intellectual history that suggested that different interpretations of

---

<sup>12</sup> In some cases, an ‘edge’ may tie two nodes that are really part of the same concept, such as ‘speech’ and ‘acts’. This appears differently from other frequently paired terms (such as ‘social’ and ‘action’) in that we would have a disconnected pair, as ‘speech’ and ‘act’ would have a much closer relation to each other than to any other terms, whereas both ‘social’ and ‘action’ would be tied to *other* concepts.

**Table A1:** Habermas Chapter6 semantic relations

<i>Word 1</i>	<i>Word 2</i>	<i>Frequency of relation</i>
Social	System	384
Lifeworld	Social	360
Lifeworld	Action	296
Lifeworld	System	286
Social	Action	270
Lifeworld	Situation	259
Social	Integration	207
Social	Society	198
Action	System	191
Communicative	Action	190
Situation	Action	187
Communicative	Lifeworld	177
Lifeworld	Society	176
Lifeworld	Concept	171
Understanding	Action	169
Lifeworld	Knowledge	168
Differentiation	Social	162
Differentiation	System	159
Lifeworld	Understanding	155
Society	System	143
Integration	System	143
Lifeworld	Perspective	136
Lifeworld	Structures	136
World	Social	133
Validity	Action	132
Action	Society	131
Level	System	130
Lifeworld	Forms	130

the term defined the three generations of the Frankfurt school. The same basic analytic approach can also be applied to determine how an author ‘uses’ (refers to) *another* author (source). Here, we focus on the subgraph that arises when we only examine the terms that are tied to the source (again, by being placed in the same paragraph with the name of the source). In the example used in the text, the source being investigated is Freud. We do this because, again, intellectual history suggests a split in the school which turned on interpretations of Freud. Here we reported only numerical results, the basis of which we go on to explain.

### Quantifying intersections

Determining the numerical degree of overlap between two concept maps requires careful analytic choices. First, two texts can differ in their length, and in the size of their paragraphs, which means that even the same underlying ‘true’ concept map might appear differently given the procedures laid out below. Second, it can be difficult to compare the overlap in *edges* between graphs that differ in their sets

**Table A2:** Shared semantic ties among Horkheimer/Adorno’s, Habermas’s and Honneth’s reason-centered subgraphs

<i>All three</i>	<i>Horkheimer/ Adorno–Habermas</i>	<i>Habermas–Honneth</i>	<i>Horkheimer/ Adorno–Honneth</i>
Reason–Social	Reason–Social	Reason–Social	Reason–Social
Reason–Nature	Reason–Nature	Reason–Nature	Reason–Nature
Social–Nature	Social–Nature	Social–Nature	Social–Nature
		Social–Philosophy	Nature–Power
		Social–Theory	Nature–Human
		Reason–Adorno	Nature–Form
		Reason–Instrumental	Nature–Thought
		Reason–Philosophy	Power–Social
		Reason–Theory	Power–Reason
		Reason–Horkheimer	Power–Human
		Adorno–Theory	Power–Form
		Adorno–Horkheimer	Power–Thought
		Philosophy–Theory	Social–Human
		Theory–Horkheimer	Social–Form
			Social–Thought
			Reason–Human
			Reason–Form
			Reason–Thought
			Human–Form
			Human–Thought
			Form–Thought

of *nodes* (and, indeed, their number of nodes). Here, our analytic choices are driven by the comparisons of interest.

First, we separate an analysis of the overlap in nodes from a second one, which is the overlap of edges *conditional* on shared nodes. That is, we restrict ourselves to the set of shared nodes. As it is possible to share nodes but not edges, though the reverse is impossible, this allows us to determine the nature of the overlap as well as its extent. Do authors use the same terms, but connect them differently, or do they make the same connections among these terms? Second, rather than choose a somewhat arbitrary standardization procedure for the quantification of overlap given different sizes of node sets – there are different formulae used for such comparisons and none are perfect – we here force the node sizes to be (nearly) identical by selecting the 100 most commonly used terms for each.

We make this selection by choosing all nodes that appear at least some number of times; however, there may be a large number of ties precisely at this cut-off value, so that including them puts us over our target of 100, and excluding them puts us under. Rather than arbitrarily drop some nodes to reach 100, we attempt to come as close to 100 as possible without making any arbitrary decisions; thus the actual number of nodes included in each graph is not always 100. For Adorno, we only use 73 nodes, but this is inconsequential, in that his near-zero degree of overlap with *any* of the other writers incontrovertibly places him at an



extreme. For the other works, the number of nodes varies from 95 (Marcuse's *Eros and Civilization*) to 108 (Fromm's *The Sane Society*). We then divide the number of shared nodes by whichever node set was smaller; thus if we were comparing *Eros and Civilization* with *The Sane Society* we would choose 95 as the denominator, since the number of shared nodes could not be larger than this. Thus, all three of Fromm's works share 29 nodes of their top 100 (or top 108 in the case of *The Sane Society*) with Marcuse's *Eros and Civilization*, leading to an overlap score of  $29/\min [95, 108] = 0.31$ , which is basically the same magnitude as the intersection *within* Fromm's three works (which varies from 0.28 to 0.33). This suggests a very strong thematic agreement between Fromm and Marcuse as to what 'Freud' means for critical theory.

We then also measure the proportion of shared *edges* within these shared nodes. Here, rather than count as 'agreement' when two maps both *fail* to include a possible edge, we simply take the number of nodes that are in *both* of our two texts' top 100, and then see how many edges connecting these nodes are present in *either* work, as well as paying attention to how many of these edges are shared. We thereby produce a number that varies from 0 (no overlap in connection made between shared concepts) to 1 (perfect overlap in connections made between shared concepts). For example, between the 29 nodes held in common by *Eros and Civilization* and *Escape From Freedom*, there are 55 connections made by Fromm, and 53 made by Marcuse. Of these, nine are shared, and thus we quantify the overlap as the number of shared edges by the number of total edges, or  $9/(55+53-9) = 0.09$ .<sup>13</sup> Given that this number ranges from 0.06 to 0.13 for *within* Fromm's own works, the overlap with Marcuse of 0.29 in Fromm's third book strongly suggests that he is doggedly pursuing similar themes.

---

<sup>13</sup> While this procedure does not correct for the density, which affects the likelihood that two edges would overlap 'by chance', the density of these graphs is so low (an average of 0.105) that the more interpretable direct approach taken here seemed preferable than the construction of a probability distribution.