THEORY AND SOCIAL INQUIRY

Brubaker, Rogers, Marion Fourcade, Kieran Healy, Greta R. Krippner, and Daniel Hirschman. 2025. "Comment and Reply: 'The Person of the Category'." *Theory and Social Inquiry* 1 (2): 1–39. DOI: https://doi.org/10.16995/tsi.26284

Open Library of Humanities

Comment and Reply: "The Person of the Category"

Rogers Brubaker, University of California, Los Angeles, brubaker@soc.ucla.edu
Marion Fourcade, University of California, Berkeley, fourcade@berkeley.edu
Kieran Healy, Duke University, kieran.healy@duke.edu
Greta R. Krippner, University of Michigan, krippner@umich.edu
Daniel Hirschman, Cornell University, dan.hirschman@cornell.edu

Editorial Note: In 2023, Rogers Brubaker submitted a comment to *Theory and Society* on an article that Greta Krippner and Daniel Hirschman had published the year prior in the same journal. Richard Swedberg, an editor at *Theory and Society* at the time, accepted Brubaker's comment and decided to also solicit an additional commentary from Marion Fourcade and Kieran Healy. Before Fourcade and Healy had an opportunity to produce their comment, a conflict between *Theory and Society*'s commercial publisher and the editorial board resulted in the resignation of the full editorial team of the journal, delaying the completion of this project. We are delighted to now publish the full exchange in *Theory and Social Inquiry*, including commentaries by Brubaker and Fourcade and Healy, as well as a reply by Krippner and Hirschman.

In "The Person of the Category" (*Theory and Society*, September 2022), Krippner and Hirschman explore how the proliferation of scoring technologies (popularly known as "algorithms") may be changing the landscape of political contestation. They develop their analysis by contrasting two paradigmatic classification technologies used in pricing decisions in insurance and credit markets: one organized around what they term "class-based" techniques typical of the pre-digital world, and another using "attribute-based" methods associated with the advent of digital technologies. Krippner and Hirschman argue that the key distinction between these two classification technologies is how they form groups and thereby enable (or disable) collective action. Class-based decision technologies form groups in which each individual holds select characteristics in common with others, allowing for the identification of shared conditions and enabling action in concert. By contrast, attribute-based techniques evaluate each individual on a series of attributes that need not be shared with others. Krippner and Hirschman argue that the latter classification technology is inherently depoliticizing, as it makes the circumstances under which individuals attach themselves to (sociologically relevant) groups more tenuous.

The Declining Significance of Classification

Rogers Brubaker

Abstract

Scoring and classification systems are ordering devices that make the social world legible and tractable. Both mediate between difference and inequality. But scoring systems organize and represent difference in gradational rather than categorical terms, through measurement rather than classification. This alters the relation between difference and inequality. Who *gets* what is linked not to who *is* what, as inferred from stable and publicly legible categorical identities, but to who *does* what, as recorded in variable and machine-readable behavioral traces. Scoring renders difference as performance; classification renders difference as identity. This comment therefore argues that individualized scoring algorithms and, more generally, new calculative technologies enabled by machine learning point to the declining significance of classification.

Greta Krippner and Daniel Hirschman's "The Person of the Category: The Pricing of Risk and the Politics of Classification in Insurance and Credit" (2022) addresses crucially important questions about the ways in which the constitution of both groups and individuals—as well as the relationship between them—is mediated by changing technologies of classification, quantification, and analysis. The article is conceptually sharp, historically rich, empirically grounded, and theoretically illuminating. The wonderfully rich footnotes alone are worth the price of admission. The article represents the old *Theory and Society* at its best.

My comment is prompted by engagement rather than by disagreement. I find compelling both the concrete comparative historical argument about insurance pricing and credit scoring and the more abstract argument about the social and political implications of class-based and attribute-based systems. What I want to propose is therefore not a critique but a way of extending—and in part reframing—the paper's concluding discussion of the broader significance of changing calculative technologies.

Krippner and Hirschman situate their argument on the terrain of the politics of classification, and they analyze the intellectual technologies developed and deployed in the credit and insurance industries as technologies of classification. But is the notion of classification perhaps being stretched too far? If "calculations of individualized scores" represent the "(digital) future" rather than the "(analog) past" (714), and if such scores entail the "dissolution of aggregates" and potentially the "dissolution of the individual as acting subject" (715), is it then helpful to conceive of such individualized scoring as a particular kind of “classification system” (714)? More generally, do artificial intelligence and machine learning, insofar as they “accelerate and amplify the trends we associate here with the early development of credit scoring,” represent the “future of the politics of classification” (714n)? Or might we instead understand the increased prominence of individualized scoring—and, by extension, of certain other calculative technologies used in algorithmic governance and supported by machine learning—as *diminishing* the significance of classification?

Classification requires classes. As Krippner and Hirschman’s paper clearly shows, the attribute-based systems that emerged in credit scoring, unlike the class-based systems used in the pricing of insurance, do not generate classes. The attribute-based logic enabled by multivariate statistics (and now by machine learning) generates scores, not classes. As the number of attributes taken into account in generating scores increases, Krippner and Hirschman suggest, attribute-based systems tend toward assigning each individual a unique score. Thus attribute-based systems do not classify individuals; they score them. Scoring, I would argue, is not a *form* of classification; it is an *alternative* to classification.

This might seem like semantic nitpicking, a distinction without a difference. Both scoring and classification, after all, are devices of ordering, ways of making the social world at once legible and tractable. In Bourdieusian language, both are principles of vision and division of the social world; both powerfully shape subjectivities; and both, as I will argue below, link difference and inequality. But they work in different ways and produce different consequences, for reasons well specified by Krippner and Hirschman. We know from a rich sociological, anthropological, and historical literature that classifications often become focal points of organization and sites of contestation and resistance; as Krippner and Hirschman remind us, they are at least potentially *collectivizing* technologies. Scoring systems, by contrast, “detach individuals from groups…altogether” (709); they are powerfully *individualizing*, even atomizing technologies (719), fostering an ethos of self-entrepreneurship (Fourcade and Healy 2017, 19–20; Brubaker 2020, 33–39).

Scoring and classes are, of course, not mutually exclusive. Scores may be grouped into classes and communicated in categorical rather than—or in addition to—numerical form. FICO groups its credit scores, for example, which range from 300 to 850, into five categories, labeled exceptional, very good, good, fair, and poor. But these categories are interpretative aids to the consumer, not operative categories for creditors. They allow consumers to know roughly where their score falls within the distribution of scores, but the thresholds between them have no general significance for the granting or denying of credit. Creditors do use cutoff values in determining the particular forms of credit to offer and at what prices, but these vary from creditor to creditor and from project to project (FICO 2023; FICO n.d.).

Krippner and Hirschman observe that creditors, like insurers, are "in the business of classification" (706). But in what sense is this the case? To be sure, creditors must draw distinctions and make "decisions about who should receive credit and what interest rate they should be charged" (706). And there is a sense in which such decision-informing distinctions entail classification: each distinction divides some population into at least two classes (in the purely logical sense of "class") whose members are to be treated differently.

These local acts of line-drawing, however, are what one might call *nonce-classifications*: they create single-use, narrowly context-bound classes. Such nonce-classifications are far removed from the enduring, consequential, and massively cumulative *systems* of social classification on which the historical and social scientific literature on classification has focused. As the latter have changed—in the domains of sex, gender, and sexuality, for example; or race, ethnicity, and nationality; or medicine and psychiatry—the "space of possibilities for personhood" (Hacking 1986, 229) and the space of possibilities for group formation (as Krippner and Hirschman note) have changed as well. These systems of classification have been deeply implicated in the structures and processes of durable "categorical inequality" that link categorical difference to inequality, who *is* what to who *gets* what (Tilly 1998; Brubaker 2015). For that reason, they have been the focal points of what Bourdieu calls "classification struggles."

Even if the nonce-classifications generated by particular lenders' use of particular credit score thresholds are relatively inconsequential, the scoring *systems* used not only in the credit industry but in an increasingly wide range of domains are hugely consequential. And they certainly contribute to durable, cumulative forms of inequality by linking difference and inequality (Fourcade and Healy 2013). But systems of scoring construe difference in a fundamentally different way than systems of classification do: they organize and represent difference in *gradational* rather than *categorical* terms. Difference is constituted, then, through measurement rather than classification. This fundamentally alters the relation between difference and inequality. Who *gets* what is linked not to who *is* what, as inferred from stable and publicly legible categorical identities, but to who *does* what, as recorded in variable and machine-readable behavioral traces. Scoring renders difference as performance; classification renders difference as identity.

I am therefore reluctant to subsume scoring under the overarching rubric of classification as Krippner and Hirschman do, following an influential line of analysis developed by Marion Fourcade and Kieran Healy. For Fourcade and Healy (2013, 560), credit scoring techniques generate "classification situations"—a concept they propose by analogy with Weber's "class situations"—that have "distinctive and class-like effects on life-chances and social identities." Their analysis of the new forms of stratification engendered by the widespread use of credit scoring technologies is brilliantly illuminating, but it leaves the notion of classification unspecified. Fourcade and Healy characterize market institutions as "inveterate classifiers." But their gloss on this—"they count, rank, measure, tag, and score on various metrics" (ibid, 562)—suggests that classification figures in their argument as an extremely general umbrella concept for *all* ways of organizing and apprehending the social world in a differentiating manner, gradational as well as categorical.[1] It is a fruitful move to bring into focus distinctive ways of "seeing like a market" (Fourcade and Healy 2017). But classification in a more specific sense—involving the establishment of a set of categories or classes and the assignment of instances to one category or another—is arguably much less central to seeing like a market (or to seeing like an algorithm) than it is (or was, in the age of "high modernity") to seeing like a state.[2] The assimilation of scoring to classification blurs the difference between gradational and categorical ways of ordering the world and making it legible, calculable, and tractable.

Scoring is regularly paired with classification, the calculation of numbers with the drawing of lines. In the credit industry, as I suggested above, this happens when a creditor offers credit at a particular price to those scoring above a particular threshold on a particular metric. But even when they are thus paired—even when a continuous distribution of scores is segmented into discontinuous categories—the scoring and the classification remain distinct moments, produced by different intellectual technologies and undertaken by different actors. A "consumer reporting agency" or CRA (generally Experian, Trans Union, or Equifax) calculates the scores by applying a proprietary scoring formula (provided most often by FICO) to regularly updated data furnished by credit-card companies, mortgage lenders, and others.[3] A lender (or landlord, insurance company, utility company, etc.) specifies the threshold and makes decisions accordingly,

[1] See also Fourcade 2016 on "phantom de-categorization," which raises issues beyond the scope of this comment (186–188).

[2] On "high modernism" as an ideology emphasizing top-down planning and the "rational design of social order," see Scott 1998, 4. As states have more recently adopted data-mining, behavior-tracking, and algorithmically mediated modes of governance, the differences between seeing like a state and seeing like a market have diminished. Yet states, governing through the classificatory medium of law, remain engines of classification in a much more robust sense than markets. On changing modes of state governance in the age of big data, see Fourcade and Gordon 2020.

[3] Since each of the three major credit agencies has its own, slightly different database of information, the same person's FICO score may differ slightly depending on which credit agency reports the score.

accessing the scores from one or more CRAs. Neither FICO nor the CRAs specify thresholds (except, as noted above, as rough consumer-facing guidelines as to where particular bands of scores fall in the overall distribution of scores). The scoring agencies are in the business of differentiation, but they are not in the business of classification; they do not establish classes. The creditors who allocate credit using score cutoffs do generate classes (in the minimal logical sense discussed above) of eligibles and ineligibles, yet the thresholds—and therefore the classes—vary from creditor to creditor and from configuration to configuration for a given creditor.[4] This works against the reification of particular thresholds and the solidification of sets of people with credit scores within particular ranges into durable and recognizable classes (Rieder 2016, 49).[5]

In the closing pages of their article, Krippner and Hirschman suggest—correctly, in my view—that big data and machine learning are amplifying the scrambling of social identities introduced by attribute-based scoring systems (717).[6] This arguably means that the burgeoning field of algorithmic governance as a whole—not just credit scoring—is fundamentally built upon gradational and individualizing rather than categorical and collectivizing ways of knowing and acting on the social world. Of course, lines must still be drawn, and thresholds specified.[7] But to the extent that machine learning and superabundant behavioral data allow people to be apprehended, tracked, identified, targeted, nudged, and governed as individualized, continuously updated data-bundles rather than as stably identifiable members of enduring social categories, then those

[4] The minimum credit scores required by Fannie Mae for manually underwritten loans, for example, vary depending on whether the loan is for a purchase or a cash-out refinance, on the number of units in the property, on the borrower's debt-to-income ratio, on the loan-to-value ratio, and on whether the mortgage is fixed rate or adjustable. This yields six different credit score thresholds for Fannie Mae loans, ranging from 620 to 720 (Fannie Mae 2025).

[5] A similar point about the distinction between a continuous distribution of numbers and the moment of line-drawing can be made about predicted probabilities, another calculative technology in which machine learning is increasingly implicated. The predicted probability that some event of interest might occur—that a defendant released on bail will flee or commit a crime; that a cancer will recur—is central to decision-making in criminal justice, medicine, and many other fields. Such predicted probabilities may be aggregated into classes, but that aggregation is a separate moment; the predicted probabilities themselves—like scores—are not classes. For debates about presenting scores in numerical or categorical form, see Scurich 2018 (on risk assessment instruments in criminal justice) and Wynants et al. 2019 (on risk prediction models in medicine).

[6] On machine learning and the scrambling of social identities, see also John-Mathews and Cardon (2022).

[7] Many machine learning algorithms are considered "classifiers" in the technical literature. But "classification" means something very specific in this literature, quite different from its meaning in the broad social science literature on the politics of classification. A spam-detection algorithm is a "classifier" in the technical sense: having been trained on data that humans have labeled "spam" or "not spam," the algorithm's job is to predict which of these two classes an incoming email falls into and to route the email appropriately. But "classifying" an email as spam or not spam is not an act of *classification* in a robust sociological sense; it is an act of *identification* or *recognition*. There is a single focal category, not a system of categories; the task is to identify a new instance as X or not X. At most it generates what I have called "nonce-classifications." (There are, to be sure, other ways in which machine learning algorithms are involved in classification, but they are beyond the scope of this discussion.)

pre-existing categories lose much of their operative significance[8] and classification (in the strong sense, pertaining to enduring classification systems of broad, cross-domain reach) becomes less important as a technology of governance.

This accelerating shift, as Krippner and Hirschman argue, has profound implications for group formation, the shaping of subjectivity, and the possibilities for collective action. It also has profound implications, as I suggested above, for the relation between difference and inequality. The mechanisms that mediate between social differentiation and unequal outcomes—and the ways in which social differentiation is made legible and actionable—are increasingly gradational rather than categorical.[9] They depend on an infrastructure of quantification, commensuration, datafication, ubiquitous tracking, and machine learning, rather than on systems of social classification. Credit scores, with their proliferating off-label uses, are one key example of the gradational linkage between difference and inequality (Rona-Tas 2017). But so are the algorithmic systems used in the hiring, assessment, and termination of workers; in bail and sentencing decisions and in educational settings; and in the generalized metrics that purport to capture civic trustworthiness or civic merit emerging in China, known as the "social credit system" but going well beyond the domain of creditworthiness.[10]

Systems of algorithmic governance, like earlier procedures built on notions of "mechanical objectivity," have been promoted in part for their putative blindness to the categorical distinctions on which formal or informal systems of categorical exclusion or differential treatment have rested.[11] These claims to procedurally guaranteed impartiality and objectivity have been widely and legitimately criticized; a large literature addresses various forms of algorithmic bias and pursues strategies for reducing or overcoming such bias.[12] I do not wish to enter into these debates. What I want to underscore is that even a perfectly impartial system of algorithmic governance,

[8] Even if categories lose much of their *operative* significance, in the sense that people are not directly governed *as* members of categories, they do not lose their broader *social and political* significance: a history of categorically organized and governed inequality, for example, remains encoded in the present, even if present mechanisms of scoring and allocation take no account of categories. I take it that this is what Rieder meant when he wrote that "The problem, here, is not that data mining can be biased, but that, after centuries of inequality and discrimination, empirical reality is biased" (2016, 50).

[9] For a critical appraisal of Tilly's (1998) account of categorical inequality, arguing that inequality has become less categorical even as it has become more severe, see Brubaker 2015, Chapter 1.

[10] On "automated hiring platforms," see Ajunwa and Greene 2019; on algorithmic management of labor, Kellogg et al. 2020. On algorithmic criminal justice, see Huq 2019. On predictive analytics in universities' efforts to model undergraduate student behavior and nudge students in particular directions, see Whitman 2020. On the social credit system, see Liu 2019. On the forces pointing toward the emergence of a generalized metric that would function as a kind of "Übercapital," see Fourcade and Healy 2016.

[11] For an historical account of objectivity and its association with impartiality and fairness, see Porter 1995 (3–8, 213–6).

[12] The problem is complicated by the fact that there are always multiple ways of construing bias, and there is no single objective standard against which bias can be measured. See for example Huq 2019 (1054ff) and Kleinberg et al. 2019.

however defined, would not reduce the *degree* of inequality; it would simply change the *form* of the relation between difference and inequality.

The shift from categorical to gradational ways of construing difference may even *intensify* certain forms of inequality by enabling powerful, data-rich firms or states to capture additional value or tighten webs of control through finer-grained discriminations. As Krippner and Hirschman note, it is often claimed that data correlated with basic social categories like race or gender serve as proxies for those categories (Harcourt 2015). But from the perspective of a profit-maximizing firm, a data-hungry state, or any organization interested in prediction (in the domains of medicine, criminal justice, or child welfare, for example), categories have long served as proxies for previously unavailable quantitative data. Categories are lossy approximations; the shift from categorical to gradational data can therefore facilitate a *tighter mapping of difference onto inequality*. Extrapolating certain contemporary trends into the realm of social science fiction—yet a fiction close enough to our own realities to be worth paying attention to—one can envision a post-categorical "meritocratic" dystopia in which rewards and punishments, opportunities and restrictions are meted out in precise proportion to scores and predicted probabilities calculated from behavioral trace data. Such a "meritocracy" would be made all the more dystopian not only by the error- and bias-ridden nature of systems of algorithmic governance and the data upon which they rely, but perhaps even more by the predictive accuracy of such systems.[13] The declining significance of classification in an age of algorithmic governance, then, is at best an ambivalent phenomenon, not an emancipatory development to be celebrated.

[13] It is often forgotten that Michael Young's 1958 fable *The Rise of the Meritocracy* was a critique of meritocracy, not a celebration. On big data and "meritocracy" as a normatively problematic ideal, see especially Rieder 2016, 51–52: "The idea that Big Data 'works'... may, then, be much more terrifying than its possible failure. ...What if the problem is knowing individuals and groups too well rather than not well enough? ... What if the saying 'ignorance is bliss' holds true for society more generally, in the sense that *not-knowing* create spaces where a plurality of practices and lives is possible because we cannot mechanically relate them to notions of performance and profit?"

Competing Interests

The author has no competing interests to declare.

Sources Cited

Ajunwa, Ifeoma, and Daniel Greene. 2019. “Platforms at Work: Automated Hiring Platforms and Other New Intermediaries in the Organization of Work.” In *Work and Labor in the Digital Age*. Vol. 33, *Research in the Sociology of Work*, edited by S. P. Vallas and A. Kovalainen. Emerald Publishing Limited.

Brubaker, Rogers. 2015. *Grounds for Difference*. Harvard University Press.

Brubaker, Rogers. 2020. “Digital Hyperconnectivity and the Self.” *Theory and Society*.

Fannie Mae. 2025. “Eligibility Matrix.” https://singlefamily.fanniemae.com/media/20786/display.

FICO. 2023. “Understanding Fico Scores.” https://www.myfico.com/credit-education-static/doc/education/Understanding_FICO_Scores_5181BK.pdf.

FICO. n.d. “What Is a FICO Score?” https://www.myfico.com/credit-education/what-is-a-fico-score.

Fourcade, Marion. 2016. “Ordinalization: Lewis A. Coser Memorial Award for Theoretical Agenda Setting 2014.” *Sociological Theory* 34(3): 175–95.

Fourcade, Marion, and Jeffrey Gordon. 2020. “Learning Like a State: Statecraft in the Digital Age.” *Journal of Law and Political Economy* 1(1).

Fourcade, Marion, and Kieran Healy. 2013. “Classification Situations: Life-Chances in the Neoliberal Era.” *Accounting, Organizations and Society* 38(8): 559–72.

Fourcade, Marion, and Kieran Healy. 2017. “Seeing like a Market.” *Socio-Economic Review* 15(1): 9–29.

Hacking, Ian. 1986. “Making Up People.” In *Reconstructing Individualism: Autonomy, Individuality, and the Self in Western Thought*, edited by T. C. Heller and C. Brooke-Rose. Stanford University Press.

Harcourt, Bernard E. 2015. “Risk as a Proxy for Race: The Dangers of Risk Assessment.” *Federal Sentencing Reporter* 27(4): 237–243.

Huq, Aziz Z. 2019. “Racial Equity in Algorithmic Criminal Justice.” *Duke Law Journal* 68: 1043–1144.

John-Mathews, Jean-Marie, and Dominique Cardon. 2022. “The Crisis of Social Categories in the Age of AI.” *Sociologica* 16(3): 5–16.

Kellogg, Katherine C., Melissa A. Valentine, and Angéle Christin. 2020. “Algorithms at Work: The New Contested Terrain of Control.” *Academy of Management Annals* 14(1): 366–410.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Cass R. Sunstein. 2019. “Discrimination in the Age of Algorithms.” *Journal of Legal Analysis* 10: 113–74.

Krippner, Greta R., and Daniel Hirschman. 2022. “The Person of the Category: The Pricing of Risk and the Politics of Classification in Insurance and Credit.” *Theory and Society* 51(5): 685–727.

Liu, Chuncheng. 2019. “Multiple Social Credit Systems in China.” *Economic Sociology: The European Electronic Newsletter* 21(1): 22–32.

Porter, Theodore M. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton University Press.

Rieder, Bernhard. 2016. “Big Data and the Paradox of Diversity.” *Digital Culture & Society* 2(2): 39–54.

Rona-Tas, Akos. 2017. “The Off-Label Use of Consumer Credit Ratings.” *Historical Social Research / Historische Sozialforschung* 42(1(159)): 52–76.

Scott, James C. 1998. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Yale University Press.

Scurich, Nicholas. 2018. “The Case against Categorical Risk Estimates.” *Behavioral Sciences & the Law* 36(5): 554–64.

Tilly, Charles. 1998. *Durable Inequality*. University of California Press.

Whitman, Madisson. 2020. “‘We Called That a Behavior’: The Making of Institutional Data.” *Big Data & Society* 7(1).

Wynants, Laure, Maarten van Smeden, David J. McLernon, Dirk Timmerman, Ewout W. Steyerberg, and Ben Van Calster. 2019. “Three Myths about Risk Thresholds for Prediction Models.” *BMC Medicine* 17: 192.

Young, Michael Dunlop. 1958. *The Rise of the Meritocracy, 1870–2033; An Essay on Education and Equality*. Thames and Hudson.

The Duality of Categories and Attributes

Marion Fourcade and Kieran Healy

Abstract

Krippner and Hirschman's "The Person of the Category" presents rich historical evidence to support an argument with which we largely agree. We suggest they can simplify the task before them, however, by not pre-judging the possibilities for mobilization under attribute-based systems and by recalling that categories and attributes are dual to one another. It is the continuous, pervasive, and above all plausibly behavioral character of data-collection that matters most to modern classification and scoring systems, rather than whether the information is statistically represented in category- or attribute-based form. We discuss the character of politics in systems of this kind, emphasizing the continued salience of social categories within such systems, the tendency of attribute-based methods to generate new political fault lines, and the role of regulation in governing the scope and limits of broadly "algorithmic" social control.

In "The Person of the Category," Greta Krippner and Daniel Hirschman (2022, henceforth K&H) examine how algorithms work as technologies of classification. They contrast two empirical cases, insurance pricing and credit scoring, which they argue "present paradigmatic features of pre-digital and digital forms of classification, respectively" (K&H, 688). Their analysis is careful, and they are fully aware of how, historically, each case is neither a pure instance of its type nor historically insulated from the other. Still, they argue that the cases represent two meaningfully different and politically consequential ways of measuring and classifying people. In their view, insurance pricing is *class-based*. It works by "assigning individuals to membership in a group" (K&H, 689). This actuarial approach puts individuals in multi-dimensional categories like "Young, rural, woman driver". Credit scoring, by contrast, is *attribute-based*. It works by "considering an individual's values on a series of variables" (K&H, 689) and results in a score of some sort that attaches to them as an individual. "The critical distinction between these two systems of classification" they argue,

> is that in class-based systems each individual is given the *same value* as all other members of the group to which she is assigned, whereas attribute systems attempt to give each individual her own *unique value* or score, as closely as this can be approximated (K&H, 689).

For K&H, this matters because, at root, a class-based system is easier for people to mobilize around for the purposes of collective action, whereas an attribute-based system is atomizing. The former kinds of categories—associated with a "pre-digital" era—are, in their terms, broadly "actuarial." Actuarial categories are legible. Their components already exist out in the world. Identifying oneself through them is subjectively straightforward. For instance, "I am a young woman living in a rural area". They also offer a basis for action: "People like me should not be discriminated against". Attribute-based systems, meanwhile, are associated with the digital era. They are broadly "statistical" (though K&H do not use this term). This makes for much less in the way of actionable material for either self-identification or collective action ("My score is 627").

K&H flesh out this idea with a comparative analysis of the insurance-pricing and credit-scoring cases. They make a plausible argument. It is true that, in general, a socially entrenched category such as "Unmarried Woman" is easier to identify with as a matter of subjective experience than a detached number like "627". It is also easier to mobilize around the former than the latter when it comes to any kind of collective or political aim.

The remainder of this comment focuses on two linked points and their consequences. The first is that, as a general matter, the degree to which particular social categories are entrenched is a variable rather than a constant (Starr 1992). Category-based systems can produce empty or useless categories; attribute-based systems may throw up more opportunities for new identities than we might expect. The second is that, despite being different in their immediate presentation and often materially different in their immediate consequences, category- and attribute-based ways of presenting data are dual to one another (Breiger 1974). That is, they are aspects of the same thing, not fundamentally different ways of representing the world. If you have one, you can get to the other. While attribute-based systems are often more opaque than class-based ones in the way that K&H describe, they too have the potential to generate categories that may become bases for social identities or political action. Insofar as modern systems militate against group-formation or collective action in the way K&H have in mind, however, their "attribute-based" character is not to blame. The high-volume, large-scale, real-time and above all *behavioral* character of the data in these systems is far more consequential.

The Character of Categories

It is helpful to begin with a difficulty that K&H address early on in their article. They argue that "actuarial", class-based systems are more amenable to group formation than attribute-based ones, especially when the classes are already socially salient. However, in order to make their argument, they must first (K&H, 690–694) clear the ground by dealing with an older but influential view about the insurance industry that claims actuarial systems cannot do this. Jonathan Simon's critique of "actuarialism" developed the idea that the categories created by insurance companies and their ilk were "singularly sterile" (Simon 1988, 789); i.e., that they were "alive only in the imagination of the actuary who calculates and tabulates and not in any lived form of human association" (K&H, 691). This is awkward for K&H's argument because, in their view, it is actuarial categories that allow for the identification of (and with) the "person in the category." They want the *attribute*-based approach to be the sterile one.

K&H make a three-pronged response to Simon. First, they argue that many actuarial classifications draw on pre-existing and already chronically entrenched categories like "gender", thus making it difficult to fully "sterilize" or "de-moralize" them. Second, they claim that even the more abstract or "artificial" categories in insurance tables "are still *potential* collectivities that can under particular circumstances be activated" (K&H, 693). And third, they assert that it is the more modern, attribute-based systems that provide a much better target for Simon's original critique. Not only do such systems treat individuals as collections of attributes, they also routinely change which attributes are deemed relevant to their scores: "scoring technologies continually swap out predictors... such that there is no stable basis for constructing group membership" (K&H, 693).

Of these three points, the first seems uncontroversially true. To the extent that actuarial classifications incorporate existing categories that are not just recognizable but in fact central to social life and individual experience, they have more than enough raw material for future mobilization. The other two points are less convincing. Or rather, they are not necessary for the overall argument of the paper to succeed. The second point opens the door to the idea that there is variability in the degree to which even seemingly obscure or unlikely categories can be targets of mobilization or, in the longer term, form a basis for identity-formation in the manner of Ian Hacking's process of "making up people" (Hacking 1986). K&H acknowledge this point (690–691) in the context of Simon's effort to distinguish between "pre-actuarial" and "actuarial" classifications. Simon thought the former, though often pernicious in motive, *did* allow for the "space of possibilities for personhood" to expand, whereas the latter did not. K&H reject this distinction with respect to actuarial classifications but wish to re-introduce it for attribute-based ones. We suggest instead that there is no principled basis for deciding *ex ante* what kinds of classifications can and cannot form the basis of socially real or consequential categories. As K&H show, Simon was wrong to think that actuarial classifications were incapable of this. But in the same way, we contend, it is not necessary to argue that attribute-based classifications are similarly inept. It is better to be a more thoroughgoing Durkheimian on this point and say that, in principle, almost anything can be grist for the mill of human difference-making and boundary-drawing. Of course, this does not mean that every last feature or attribute will in fact seed some widespread social identity or successful political mobilization. But there is no point in tying our hands about it in advance.

K&H's third point against Simon brings the central issue into view. They say that the fragmentation of subjects Simon complains about is "more fully realized" by attribute-based systems not only because they generate scores rather than groups, but because the basis for calculating those scores is continually shuffled around or swapped out. As we argue in Fourcade and Healy (2024; see also Fourcade 2021), not only do modern systems routinely readjust their scoring methods, thereby putting individuals at risk of sudden reclassification for reasons that will likely be opaque to them, those methods and the classification-situations they create vary across markets. The "same" constellation of attributes can produce different kinds of experience in different market settings. However, the shuffling and swapping-out of attributes is mostly enabled by advances in information technology. It is not an intrinsic feature of an attribute-based approach. If early insurance companies had been able to easily shuffle and swap their categorical configurations, they probably would have. That they did not is more a matter of technological capacity rather than a feature of class-based methods.

Personal Data Makes Categories Dual to Attributes

We are left with the question of whether attribute-based systems really do have the features that K&H describe. These systems represent people as clusters of attributes and generally use those attributes to calculate individualized scores. Those scores really are, to begin with, less conventionally "accessible" or "legible" than a group-based classification. Nevertheless, we should not reify the distinction between attribute-based and class-based methods. They are dual to each other in much the same sense that persons and groups are dual to one another. The classic statement of this idea, which goes back to Georg Simmel, is Breiger (1974).[14] The insurance classes and categories K&H discuss are just aggregates of individual characteristics, or attributes. By the same token, scores attributed to individuals can be characterized as the result of membership in cross-classified categories.

This is perhaps easiest to show by way of example. Conveniently, and not coincidentally, the long tradition of studying and modeling the relationship between individual attributes and cross-classified tables of counts often draws on examples from the insurance industry.

[14] For a recent exploration of the notion focused on its methodological implications, see Schoon, Melamed, and Breiger (2024).

Gender	Location	Seatbelt	Injury: No	Injury: Yes
Female	Urban	No	7287	996
Female	Urban	Yes	11587	759
Female	Rural	No	3246	973
Female	Rural	Yes	6134	757
Male	Urban	No	10381	812
Male	Urban	Yes	10969	380
Male	Rural	No	6123	1084
Male	Rural	Yes	6693	513

Table 1: Example car accident data from Agresti (2012).

Consider **Table 1**, which presents counts of car accidents occurring across a number of categories of person. The data are taken from the sample datasets accompanying Agresti (2012). For our purposes, the details about the data and its limits are less relevant than its form. The table presents a simple cross-classification of categories that is directly analogous to the more complex example shown in K&H’s **Figure 1** (reproduced on page 30 below), in which cells in the insurance pricing table are labeled with codes designating what sort of person is in each group. For example, code 8291 corresponds to “Unmarried female, with driver training, aged 20, driving for pleasure use or farm use”, all within the broader category of “Youthful operator, not eligible for Good Student Credit”. Each cell is associated with a “factor” (e.g., 1.20) that corresponds to some adjustment to the insurance premium charged to people falling into this cell. This premium, for its part, would be constructed from a table of counts of accident rates like our **Table 1**. The more accidents that people in a particular cell are involved in, the higher their insurance premium will be.

Gender	Location	Seatbelt	Factor
Female	Urban	No	0.137
Female	Urban	Yes	0.066
Female	Rural	No	0.300
Female	Rural	Yes	0.123
Male	Urban	No	0.078
Male	Urban	Yes	0.035
Male	Rural	No	0.177
Male	Rural	Yes	0.077

Table 2: Results from a model predicting car accident injuries within each category.

To mimic K&H's **Figure 1**, we can construct a "factor" that captures the relative degree of risk for each group by fitting a model to the cells in **Table 1**. We want to capture variability in accident or injury rates. We are interested in the relative risks of car accident injuries for our different groups.[15] Again, the details are irrelevant here, but we can write a model (either a logistic regression on individual accidents or, equivalently in this case, a Poisson model of the cell counts in the table) estimating how much accident risk is associated with each group. We would end up with something like **Table 2**. Here, the column labeled "Factor" is our estimate of riskiness, which would serve as the basis for adjusting insurance premiums and is analogous to the one shown in K&H's **Figure 1**.

So, we can take a table of injury rates cross-classified by different kinds of driver and construct a risk score for each category based on differences in accident rates between groups. However, should we wish, we can equivalently represent this model as a series of attributes. Regression coefficients attaching to variables would then be summed to produce a predicted value, or score, for any given individual. In order to get a score for any combination of attributes, we tell the model what attributes we want, and we get some number out the other side. If we wanted, we could convert these coefficients into a more user-friendly look-up table of the sort shown in K&H's **Figure 2** (reproduced on page 31 below), so that a scorer would just have to add up numbers to get a risk score interpretable on a scale we would also provide.

The point is that these are just two different ways of looking at the same thing. We can move between them as we wish. If we want to emphasize the classes, we can look at the data as a table of cross-classified categories and the risk factors associated with each group. If we want to emphasize individual attributes, we can look at the same data as a series of personal scores. **Table 3** shows ten randomly selected individuals from our original data and their personalized predicted scores. (We have added a measure for Age to our model, just to show that we can extend the approach to attributes with more than two categories.) In a model like this, there will be as many unique scores as there are unique combinations of attributes across the individuals observed in our data.

A reader might object that modern attribute-based systems, like credit scoring methods, are far more complicated than this toy example; that they involve hundreds or thousands of attributes; that they are changing all the time; and that individuals under these circumstances are in no position to establish their categorical membership. These are all perfectly reasonable responses. But it remains that there is nothing *in either the*

[15] These example data are for injuries in cases where a car accident has already happened. But things would work the same way if the event was "Got in a car accident Yes/No" rather than "Injured in a car accident Yes/No". To get an individual-level table suitable for a logistic regression, we "uncount" each cell of the summary table so that we have, for instance, 996 individual observations of Urban, Female, Seatbeltless drivers who suffered an injury.

Gender	Location	Seatbelt	Age	Score (×100)
Male	Rural	No	23	17
Male	Urban	No	42	8
Male	Urban	Yes	71	3
Male	Rural	Yes	19	8
Female	Rural	Yes	78	13
Female	Rural	Yes	65	13
Female	Rural	No	41	29
Male	Urban	No	74	7
Male	Rural	Yes	52	8
Female	Urban	No	33	14

Table 3: Ten randomly selected "individuals" and their personalized scores.

data or the modeling that makes one approach an intrinsically "class-based" and the other an intrinsically "attribute-based" way of seeing the world. In a quite literal way, they are the same thing seen from two different perspectives.

Moreover, categorical representations can be complex, too. K&H's **Figure 1**, for example, shows just a small part of a table with a total of 161 cells or classes.[16] That is a complex classification that would be difficult to grasp all at once. It probably contains quite a few cells that are very sparsely populated, that is, category combinations that are not especially common or typical in terms of social identities. Indeed, this sort of complexity—the way actuarial cross-classifications tend to expand multiplicatively—is what led to Simon's critique of the sterility of actuarial schemes in the first place. By the same token, the "attribute-based" credit scorecard shown in K&H's **Figure 2** consists of numbers that add up to a score. But each of those numbers is derived from a categorical variable. That is what the table of cells *is*. Certainly, the specific score one ends up with on that scorecard does not correspond directly to a social category, in the sense that different combinations of attributes could sum up to the same score. But the same is true of the class-based table in K&H's **Figure 1**, just from the other side: Code 8392 (Married male with driver training, aged 19, drives to work) has the same score as Code 8222 (Unmarried Female, without driver training, aged 18, drives to work).

For this reason, it can be a little misleading to speak of the "unique individual score" generated by an attribute-based system and set it against a seemingly much simpler adjustment factor associated with a category system. While it is true that, in an "attribute" view, every unique individual has a score of their own, it is not true that every individual's score is therefore unique. Hundreds of millions of unique individuals

[16] K&H are of course aware of the large number of classes in the Figure; they mention it explicitly.

in the United States have a FICO credit score of their own. But a FICO score is simply an integer between 300 and 850. Millions of individuals have the same score.

In practice, even the fine-grained character of attribute systems tends to be employed in a coarsened manner. Both class- and attribute-based systems tend to make-up people (or, less dramatically, define classes of customers or identify market segments). Most obviously, in the case of credit scoring, the gradations of points-based scoring schemes—which, in the end, remain discrete rather than truly continuous—are typically collapsed into much simpler classifications ("Good/Acceptable/Bad"; "Prime/Subprime"). Those classifications look much more amenable to individual identification and political mobilization.

We see this all the time, with greater or lesser degrees of plausibility and with greater or lesser degrees of success. Categories derived from market research may become shorthand for kinds of person; classes developed by survey researchers ("Soccer Moms") loop back to become the basis for electoral competition; and simple designations like "Good Credit" or "Bad Credit" become recognizable in the market despite the complexity of the underlying scoring technologies and the heterogeneity of people thus classified. These social processes play out in both "category-based" and "attribute-based" systems, again because these are just two views of the same underlying thing. It is more important to think about what the politics of classification looks like across these systems and what it looks like in terms of the processes of data collection and application, than to focus on an invidious distinction between them.

The Demoralization of Politics or The Politics of Demoralization?

The person is what ties category and attribute together, what makes them dual to one another. Modern attribute-based credit scoring systems (and their descendants) have far more dials to turn and switches to flip than the original round of class-based actuarial methods, and they can do it all much faster. But it is a mistake to think one approach precludes the other. Again, they are dual to one another, in both method and substance. If there is a real difference between the older way and the new, it is not the distinction between class and attribute. Rather, it is the tendency—itself the outcome of both political demands and technological change—for newer systems to present themselves as collecting fundamentally *behavioral* data and to have the capacity to do this in something like real time. Much of the moralizing power of our current engines of eigencapital derives from their claim to capture what people *do* as a basis for describing the sort of people they *are*. What matters is the ability to implement a broadly behavioral conception of social data[17] and to capture and represent it as a steady flow of information

[17] Behavioral assessments have always been part of risk profiling in both insurance and credit (Carruthers 2022, Sadowski 2024), but they were generally used in combination with other characteristics, such as race (or a proxy thereof, such as

about choices and decisions, regardless of whether this flow is represented in terms of group categories or individual attributes.

The persistent salience of categories

Rather than inadvertently reifying a distinction between two types of classification, it is far more fruitful to ask how political action tends to express itself as one or other aspect is foregrounded. When “categories” are salient, they may become the basis for collective action just because of their fit with already-existing social groups. Very often, political action on these grounds is essentially a push for equal treatment, effectively an effort to make the class less relevant. Thus, when the National Organization for Women (NOW) sued Pennsylvania’s auto insurers for discrimination against low mileage drivers (who happened to be women), they made a move to elevate individual behavior over “demographics” as the legitimate way to judge a person’s insurable risk. They lost that particular suit, but in the long run, they won the war. The state’s regulators banned gender-based insurance pricing (K&H, 704), and insurers increasingly treated “miles driven” as a useful “attribute” in premium calculations. Similarly, when NOW’s Credit Task Force dropped its plan for legal action against sex discrimination in credit markets, they did so because the problem had by then “slipped from view” (K&H 2022, 713). Once creditworthiness was reduced to a bundle of behaviors, discrimination became much harder to prove. Much the same was true of racial discrimination (Norris, forthcoming).

For the purposes of calculating risk scores or insurance premiums, “Woman Driver” and “Low Mileage Driver” are both labels that can be thought, as a matter of convenience, as either individual-level attributes or group-level categories. It is not that one is intrinsically a demographic category and the other an individual-level attribute. What makes a “class” relevant to politics, in other words, is not that it can be represented as a categorical variable in some calculation, but that it is entrenched as a social identity or established as a legal category and is thus actionable (or not) on that basis. The distinction between a group identity and an individual attribute, preference, or behavior is blurred in practice. We will tend to favor the more entrenched or better-institutionalized one, whether it is at the group- or individual-level, as more natural or fundamental. But this, too, is subject to empirical variability. In domains where legal prohibitions are limited or do not exist, such as advertising, it is still very common for organizations to use gender, age, sexuality, citizenship, religion, or race to target

neighborhood), gender, marital status, religion, or citizenship. The distinctiveness of the current regime is its propensity, rooted in legal prohibitions, to eliminate such “suspect categories” from calculations of creditworthiness, combined with its ability to collect data quickly and at scale. One reason why such categories have persisted longer in insurance (but not in credit) is that private insurance in the United States is primarily regulated at the state level while credit is subject to strong federal regulations.

specific populations. In that context, groupings are taken to be the "real" entities out of which targetable markets and individual purchases originate. When these categories are not available as public self-identifications, it is also common for data-rich advertising engines to simply proxy for them via some bundle of behaviors that *can* be observed and that is correlated with the category (Cheney-Lippold 2017). In this case, membership in the supposed underlying category is a matter of predicted probability rather than something that is simply available as a datum. It is algorithmic and behavioral, rather than fixed and declarative (Mittleman 2022). Because derived or predicted categories tend to vary depending on data availability, model design, and organizational purpose, they tend to be less legible and less stable across domains. But that does not mean that they will have no politics around them.

For instance, Facebook's "multicultural affinity" measure is constructed from behavioral data (e.g., the music people have liked on the platform), and it allows advertisers to select their targets by ethnic background (e.g., African-American, Jewish, Spanish-speaking Hispanic, and so on). Advertisers think of the ethnic categories as "real" in this case, i.e. as the entrenched category that will allow them to legibly link and label many different behaviors and preferences. It may be that the people so-classified think that way, too. But this sort of thing is dicey because it is close to the legal world of protected and suspect classes. There, the reality of these categories has a different social character, especially in the context of certain markets. In 2016, Facebook came under fire when an investigation by the online magazine *ProPublica* showed that it was allowing users to select "ethnic affinity" to advertise for rental housing (Angwin and Parris 2016). In 2019, Facebook settled subsequent legal actions by civil rights groups and removed the option to select the "ethnic affinity" measure (and others, such as age and gender) for ads in so-called "sensitive" domains.

The point is straightforward. Where demographic categories are deemed suspect by the legal system—in employment, credit, housing, or insurance—attribute-based proxies for these categories remain actionable. This is one way that politics persists in a world of data-driven targeting: through *salient behavioral proxies* for protected classes. For instance, legal scholars have argued that the kind of "affinity profiling" practiced by Facebook amounts to a form of "discrimination by association" where discriminatory practices continue through their use of correlated attributes rather than direct demographic targeting (Wachter 2020).

We should not assume that the only politics worth worrying about depends on a category being flagged as suspect by the legal system. Algorithms that are non-discriminatory from the point of view of suspect categories could still discriminate against individuals because they share some other characteristic—because they have children, for instance, or live in a high-crime neighborhood or possess some seemingly arbitrary trait that the algorithm has flagged as "risky." What kinds of politics might arise there?

The moral disentanglement of attributes

In a data-rich environment, the duality of attributes and categories manifests itself through their easy convertibility. Individual-level attributes can be used to impute group-level categories, as we have just seen. Still, categories will lurk beneath the surface of attribute-based classifications that purport to exclude them (Lockhart 2023). We can dig just a little deeper into those domains where the deployment of demographically mobilizable categories (such as gender or race) is legally prohibited. In such cases, not only are people being judged as "bundles of attributes" (in K&H's vocabulary), but algorithms are often specifically designed to make demographic categories irrelevant to predicting or explaining outcomes. Behavioral models, such as credit scores, are constructed so as to eliminate the statistical effect of suspect categories. Concretely, this means that while race or gender may be associated with some behavioral measure, it will not predict outcomes even if it is added to the model. The much-decried COMPAS algorithm, for instance, does not at first glance *appear* to be racially biased with respect to group differences in actual rates of recidivism (Corbett-Davies et al. 2016). Similarly, credit outcomes do not *appear* to be racially biased once credit scores are taken into account (Norris 2022). In both cases, the attributes selected for inclusion in the model produce a score while also washing-out any direct association with race were it to be included as a predictor. But measures like credit ratings or recidivism scores will still tend, on average, to be strongly correlated with racial classifications. This can happen for reasons varying from structural racism (e.g., minority populations are more likely to be surveilled or targeted by predatory vendors) to population-wide cultural patterns (e.g., higher levels of violence, poorer credit behavior) that have historically crystallized under conditions of racial oppression and exploitation (Patterson 2000).

There are two main ways to design a predictive algorithm that captures real, and potentially stigmatizing, discrepancies between populations. One, represented by the COMPAS case, applies where individual-level data about the behavior of interest is deficient. The other applies when data is abundant, as is the case for credit scoring. At the time of the ProPublica investigation, the COMPAS score was based on factors derived from a questionnaire with one hundred and thirty-seven items filled out by defendants about their life history, their social connections, and their neighborhood environment. Race was not one of these items. But many of the variables measured were essentially ecological (e.g., whether the defendant lived in a high-crime neighborhood) or otherwise collective (whether they had a parent who had been incarcerated), rather than behavioral. It is rare to have good, detailed data about behaviors associated with outcomes like getting arrested or convicted. In its absence, the COMPAS algorithm labeled people as members of all kinds of broader categories in order to predict their individual risk for those outcomes. The result was a widely used system deployed in courts and by parole boards around the country. It was not particularly good. It performed about as well as assessments made by novices with access to just a few basic data points (Dressel and Farid 2018). When coupled with the different absolute sizes of defendant populations by race, and differences in base-rates of recidivism within those populations, the resulting scores inevitably over-classified individual Black defendants as being at a high risk of recidivism, even though race was not used in the calculation of the risk score.

However shocking the racial disparity in recidivism scores may appear, it does not offer a good basis for political challenge under current legal rules, given that it matches empirical disparities in actual group-level rates of reoffending. The critique has to come from elsewhere. One possibility is to attack the discrepancy in reoffending statistics as a product of the centuries-long over-surveillance, over-policing, and over-penalization of Black populations (Gandy 2009; Harcourt 2015; Browne 2015; Benjamin 2019). Another strategy is to confront similar biases in the predictive tool's practice of using group-level inferences (e.g., about the defendant's neighborhood, their social connections, or their pattern of residence) to assess individual-level risk. These kinds of "attributes" may be appealing because they are "easy to score," "statistically valid" (they have a demonstrated correlation to the outcome of interest), and "intuitively plausible" (Underwood 1979). Still, it seems fundamentally unfair to judge individuals as higher risk simply because of their parents' history of imprisonment, regardless of how well this factor might predict outcomes. The fact that the stakes are so high makes using these measures all the more jarring. Not only is a person's freedom on the line, but a high score can become a self-fulfilling prophecy. Keeping a person behind bars might easily create the conditions (by way of shifts in social connections, unemployability, or any number of other pathways) that favor reoffending once they are no longer incarcerated.

Similar moral doubts play out in many social domains. Barbara Kiviat (2019; 2021) finds that both insurance regulators and the public have been skeptical about the use of credit scores for auto insurance, in spite of their effectiveness in predicting risk to the insurer. In this case, a demonstrable statistical connection between a set of attributes (condensed into a credit score) and an outcome of interest (insurance risk) is not enough to support their inclusion into the predictive tool, because this inclusion seems either unfair or irrelevant to what is being predicted. More direct evidence of driving behavior, such as data obtained from telematic devices built into vehicles, appears more acceptable. This work of moral disentanglement is the second way that politics continues to matter in what K&H call "attribute-based" systems. People often have strong intuitions about which attributes ought to be used to predict outcomes, intuitions that may sustain various forms of resistance and mobilization. The politically central one is the sense that algorithmic decision-making should rely only on behavioral records (specifically those records that pertain to the outcome of interest). Group characteristics, environmental context, and seemingly irrelevant records should be banned. This is, of course, easier said than done.

In particular, the temptation to use statistical signals of all kinds in an effort to make small (but lucrative) gains in predictive accuracy remains strong. Moreover, the structure of these political battles is underdetermined. On the one hand there is a push for "fairer" measures—that is, more fine-grained, individual, behavioral data. On the other, there is a backlash against the levels of surveillance this inevitably entails, especially since it is now increasingly automated. People may even find themselves on both sides of the political problem at once. When thinking about their insurance policies, they may favor the behavioral measures their car collects, because they are good drivers. When rushing to work, they may rage at automated speed or toll cameras as a violation of civil liberties. In each case, what is experienced or labeled as "spying" or "snooping" will vary. Opportunities for moral politics abound.

The New Frontiers of Regulatory Governance

Credit scoring is a paradigmatic example of a domain in which individual-level data is abundant, and that is perhaps why it plays such a prominent role in K&H's analysis. Like other predictive systems, it aims to maximize predictive accuracy while avoiding bias arising from the use of protected characteristics like race or gender. What makes it distinct is its seemingly exclusive focus on individual behavior. Standard methods typically do not consider group-level or environmental factors. They simply look at whether and how people pay back their loans, how much credit they use relative to some set maximum, and how long they have been in the system. Even under these conditions, it is a stretch to assume that scores are *purely* behavioral. To begin with, rules about which behaviors should be measured, and how they should be aggregated, are constantly changing. During the 2008 financial crisis, banks reduced credit card limits, which mechanically reverberated into credit scores. Large pools of people saw their credit scores drop, though their behavior had not changed. (Indeed, because that behavior was in the past, it could not have changed.) At the time of writing, credit reporting agencies commonly access people's bank accounts directly to gather detailed payment data beyond credit. In the perennially optimistic eyes of the industry, these newer, more invasive scoring models are also more inclusive. They help incorporate previously invisible, and thus excluded, individuals into the financial system. They give a boost to people who, under traditional methods, would be stuck with low scores. At the same time, the sheer intrusiveness and growing opacity of these methods invites fresh political scrutiny. Do institutions wielding algorithmic scores owe their subjects an explanation of how and why they have arrived at their decisions?

The European Union has made the boldest regulatory move. The Artificial Intelligence Act of 2024 guarantees subjects of algorithmic systems a right to an explanation, including the right to challenge automated decisions. Article 50 of the Act also establishes "transparency obligations for the providers and deployers of AI systems." Finally, there is a new awareness among policymakers the world over that "algorithms" are no substitute for effective decision-making.

Most modern technical methods can demonstrate compliance with non-discrimination rules, because they are deliberately calibrated to erase traditionally "categorical" effects. But this alone should not justify their use. The substantive point is whether a chosen method results in decisions that properly serve its intended purpose. Without independent mechanisms to verify claims to effectiveness, many companies peddling "algorithmic" solutions might be selling "AI snake oil."[18] Just as regulatory agencies seek to ensure food and drug safety, a new infrastructure is emerging that aims to separate genuinely useful automated methods from those that are ineffective or harmful. This field is developing its own standards, crafting its own policies, and training its own experts. As a form of politics, *regulatory governance* may be somewhat bureaucratic and obscure. This does not make it any less real as politics, however.

The Duality of Politics

What should we make of the apparent neutralization of actionable categories in systems like credit scoring? One interpretation would style it as a victory for civil rights. Historically underserved groups have fought long battles, which Krippner herself (2017; 2024) has extensively documented, to be evaluated as individuals outside of the socially entrenched categories that, historically, directly and intentionally determined

---

[18] "As far as we know, no hiring automation company has ever published a peer-reviewed paper validating its predictive AI, or even allowed an external researcher to evaluate it. Two of the leading companies made a show of external audits: Pymetrics contracted with a leading research group from Northeastern University, and HireVue contracted a noted independent auditor. But in both cases, the researchers were allowed to analyze only whether the AI was biased with respect to race or gender, and not whether it worked." (Kapoor and Narayanan 2024, 24)

their life-chances, mostly for the worse. In that sense, the widespread implementation of behavioral scoring can be seen as a real political accomplishment, a fulfillment of a core demand of civil rights: to be judged by what one does, rather than who one is.

A second interpretation is more skeptical. Perhaps algorithmic systems are deliberately engineered to conceal ongoing discriminatory practices rather than eliminate them. In this way of looking at things, risk-scoring methods are tortured until they produce apparently non-discriminatory predictions, the better to preserve advantage for some over others. The specific attributes selected for use in models, and the relative weight given to them, are not neutral technical choices but careful decisions calibrated to maintain social hierarchies while appearing neutral and objective.

In a sense, both perspectives are correct. Attributes and categories are dual to one another, and this duality extends to their politics. The move to escape the constraints of the data’s apparent structure—whether “downwards” in the direction of more fine-grained individuality or “upwards” in the direction of a greater degree of group solidarity—is an expression of the tendency for people and their social lives to overflow whatever organizational and institutional matrix is imposed upon them (Fourcade and Healy 2024, 61–66, 260). Whether they foreground individual-level attributes or category-based properties, we should not restrict our theory of their politics to those foregrounded elements. Indeed, we should not restrict ourselves to the formal structure of the data at all. Within every entrenched identity is a collection of persons wanting to be treated as individuals. Inside every individual-level bundle of attributes is a social identity waiting to emerge and act on its own behalf. These dual struggles reflect efforts, flawed but ineradicable, to balance individual justice with collective redress. At any point in time, they may take different forms, depending on the character of legal instruments, the scope and availability of data, and the state of statistics. But they are two faces of the same political coin.

Competing Interests

The authors have no competing interests to declare.

Sources Cited

Agresti, Alan. 2012. *Categorical Data Analysis*. 3rd edition. Wiley.

Angwin, Julia, and Terry Parris, Jr. 2016. "Facebook Lets Advertisers Exclude Users by Race." *ProPublica*, October 28. https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race.

Benjamin, Ruha. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity.

Breiger, Ronald L. 1974. "The Duality of Persons and Groups." *Social Forces* 53: 181–90.

Browne, Simone. 2015. *Dark Matters: On the Surveillance of Blackness*. Illustrated edition. Duke University Press.

Carruthers, Bruce. 2022. *The Economy of Promises*. Princeton University Press.

Cheney-Lippold, John. 2017. *We Are Data*. NYU Press.

Corbett-Davies, Sam, Emma Pierson, Avi Feller, and Sharad Goel. 2016. "A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased Against Blacks. It's Actually Not That Clear." *Washington Post*, October 17. https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/.

Dressel, Julia, and Hany Farid. 2018. "The Accuracy, Fairness, and Limits of Predicting Recidivism." *Science Advances* 4(1): eaao5580. https://doi.org/10.1126/sciadv.aao5580.

Fourcade, Marion. 2021. "Ordinal Citizenship." *British Journal of Sociology* 72(2): 154–173.

Fourcade, Marion, and Kieran Healy. 2024. *The Ordinal Society*. Harvard University Press.

Gandy, Oscar H. 2009. *Coming to Terms with Chance: Engaging Rational Discrimination and Cumulative Disadvantage*. Routledge.

Hacking, Ian. 1986. "Making up People." In *Reconstructing Individualism: Autonomy, Individuality and the Self in Western Thought*, edited by Thomas C. Heller, Morton Sosna, and David E. Wellberry. Stanford University Press.

Harcourt, Bernard E. 2015. "Risk as a Proxy for Race: The Dangers of Risk Assessment." *Federal Sentencing Reporter* 27(4): 237–43.

Kapoor, Sayash, and Arvind Narayanan. 2024. *AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How To Tell the Difference*. Princeton University Press.

Kiviat, Barbara. 2019. "The Moral Limits of Predictive Practices: The Case of Credit-Based Insurance Scores." *American Sociological Review* 84 (6): 1134–58.

Kiviat, Barbara. 2021. "Which Data Fairly Differentiate? American Views on the Use of Personal Data in Two Market Settings." *Sociological Science* 8: 26–47.

Krippner, Greta R. 2017. "Democracy of Credit: Ownership and the Politics of Credit Access in Late Twentieth-Century America." *American Journal of Sociology* 123(1): 1–47.

Krippner, Greta R. 2024. “Gendered Market Devices: The Persistence of Gender Discrimination in Insurance Markets.” *American Journal of Sociology* 130(3): 595–643.

Krippner, Greta R., and Daniel Hirschman. 2022. “The Person of the Category: The Pricing of Risk and the Politics of Classification in Insurance and Credit.” *Theory and Society* 51 (5): 685–727.

Lockhart, Jeffrey W. 2023. “Gender, Sex, and the Constraints of Machine Learning Methods.” In *The Oxford Handbook of the Sociology of Machine Learning*, edited by Christian Borch and Juan Pablo Pardo-Guerra. Oxford University Press.

Mittleman, Joel. 2022. “Intersecting the Academic Gender Gap: The Education of Lesbian, Gay, and Bisexual America.” *American Sociological Review* 87(2): 303–335.

Norris, Davon. Forthcoming. “Epistemological Slights or How Racial Capitalism Gets Away with Being Racial.” In Marion Fourcade, Greta Krippner, and Sarah Quinn (eds.) *Political Economy Rebooted*. Duke University Press.

Norris, Davon. 2022. *How All Data Became Credit Data: The Logic of Scoring and the Limits to Racial Inclusion in an Algorithmic Age*. Ph.D. Thesis, The Ohio State University.

Patterson, Orlando. 2000. “Taking Culture Seriously: A Framework and an Afro-American Illustration.” In *Culture Matters: How Values Shape Human Progress*, edited by Lawrence E. Harrison and Samuel P. Huntington. Basic Books.

Sadowski, Jathan. 2024. “Total Life Insurance: Logics of Anticipatory Control and Actuarial Governance in Insurance Technology.” *Social Studies of Science* 54(2): 231–256.

Schoon, Eric W., David Melamed, and Ronald L. Breiger. 2024. *Regression Inside Out*. Strategies for Social Inquiry. Cambridge University Press.

Simon, Jonathan. 1988. “The Ideological Effects of Actuarial Practices.” *Law & Society Review* 22(4): 771–800.

Starr, Paul. 1992. “Social Categories and Claims in the Liberal State.” *Social Research* 59(2): 263–95.

Underwood, Barbara D. 1979. “Law and the Crystal Ball: Predicting Behavior with Statistical Inference and Individualized Judgment.” *The Yale Law Journal* 88(7): 1408.

Wachter, Sandra. 2020. “Affinity Profiling and Discrimination by Association in Online Behavioral Advertising.” *Berkeley Technology Law Journal* 35(2): 367–430.

The Social Life of Scores

Greta R. Krippner and Daniel Hirschman

Abstract

We are grateful for the opportunity to clarify our argument with the aid of Brubaker's and Fourcade and Healy's generative readings of "The Person of the Category." With respect to Brubaker's critique, we maintain that scoring does achieve a *kind of classification*, albeit one that is distinct from the processes of group formation explored in a now-influential body of sociological scholarship. With respect to Fourcade and Healy's argument, we defend the importance of attribute-based versus class-based systems for shaping political action by distinguishing between *cultural* and *computational* requirements of decision technologies.

It's a tremendous thrill—and a daunting challenge—to respond to comments on our paper by three scholars whose work has deeply shaped our own. Rogers Brubaker's (2005; 2015; 2016) field-defining writings on classification, especially his insistence that we should consider "groupness" as contingent and variable rather than thinking of "groups" as fixed and bounded entities, inspired our approach to how insurers and creditors constitute groups differently and why this might matter for understanding the changing contours of political action in the digital age. Similarly, we have closely followed Marion Fourcade and Kieran Healy's fruitful collaboration since they first began exploring credit scoring as producing "classification situations" (Fourcade and Healy 2013) and later broadened these insights into a general theory of "ordinalization" as a social process and even a form of society (Fourcade 2016; Fourcade and Healy 2024).

Before engaging Brubaker's and Fourcade and Healy's astute comments, we first briefly summarize the argument of our original paper for those who have not read it. Broadly, our concern in the paper is to consider how the proliferation of new scoring technologies (popularly, if confusingly, referred to as "algorithms") across decision domains in society has reshaped potentials for collective action. To get leverage on this problem, we contrast scoring technologies used in determining access to credit to older decision technologies rooted in insurance markets. Insurance pricing relies on what we refer to as a *class-based* decision technology: individuals are assigned membership in a group defined by a set of characteristics shared with others and then treated as *average* or *typical* members of that group. The fact that one is joined with others who hold a few salient characteristics in common both constitutes the group and (potentially) facilitates collective action. By contrast, credit scoring is *attribute-based*: one's creditworthiness is evaluated by tallying values on a series of variables (or attributes) deemed important for determining the likelihood that a loan will be repaid. While multiple individuals may share the same score, each would-be borrower arrives at this score through a unique combination of values on the variables scored. As such, the group constructed by the credit score lacks the sociological foundation of the insurance risk class: one does not necessarily share characteristics in common with those who achieve the same score. For this reason, we expect that collective mobilization will be more difficult to achieve where scoring technologies predominate.

To illustrate this argument, we examine the distinct histories of contestation around the use of gender as a classifier in insurance pricing and credit scoring—paradigmatic examples of class-based and attribute-based systems of pricing risk, respectively. Insurance markets in the United States witnessed sustained contestation over gender discrimination beginning in the 1980s. Our analysis shows that the class-based nature of the pricing system allowed gender difference to "show through" other variables used to price the risk of accident. As a result, challenges to gender discrimination in insurance markets persisted, even after gender was legally prohibited as a pricing variable in some states and some lines of insurance (as in the Pennsylvania auto insurance case we examine in our article). Contestation around gender discrimination in U.S. credit markets, by contrast, was short-lived, dissipating shortly after credit scoring was introduced into these markets in the mid-1970s following the passage of a federal anti-discrimination law.[19] Here we argue that the way the credit score effectively scrambles group memberships—because individuals who achieve the same score do not share characteristics in common other than the score itself—militates against sustained collective action.

In commenting on our paper, Brubaker's main point is that classing and scoring are *fundamentally different* operations, thus it is misleading to consider scoring as a form of classification at all. Fourcade and Healy critique the paper from another direction, invoking Breiger's (1974) notion of "duality" to suggest that classes and attributes are *aspects of the same thing*, "two faces of the same political coin," as they write. In this regard, Fourcade and Healy suggest that what matters for political mobilization is not whether a decision technology is organized around classes or attributes, but the increasing reliance on data that are voluminous, continually available, and behavioral in nature. Accordingly, our goal in this reply is to attempt to navigate between Brubaker's and Fourcade and Healy's differing critiques in order to clarify ambiguities in our own argument about how these two classification technologies enable and disable different kinds of politics. We will ultimately defend our position that both classing and scoring

[19] One might reasonably conclude that the dissipation of contestation was a direct result of the passage of anti-discrimination legislation and not, as we suggest, the introduction of credit scoring technologies. Against this view, we argue that credit scoring allowed discrimination to continue by other means, as feminists understood at the time and as has been observed subsequently (see Poon 2012; Krippner 2017; Norris forthcoming).

are forms of classification *in a certain sense*, while acknowledging the way in which our presentation in the paper creates unnecessary confusion around this point. We will also stand by our view that the technologies themselves—that is, their class- or attribute-based nature—*do* matter for political action, but in ways that Brubaker and Fourcade and Healy help us to articulate more sharply than in our original paper.

Let's start with Brubaker. He points to our (repeated) assertion that credit scoring aims to give each individual a unique score as meaning that scoring should *not* be treated as a classification technology. "Classification requires classes" Brubaker argues, and since credit scores do not constitute classes but place each individual in *their own cell* (or at least aspire to do this), this is at best a single-use "nonce-classification," a logical operation and not a sociological one. This argument is laid out with precision and care, and on its own terms it is airtight. But this argument hinges on what exactly it means to give each individual a "unique score," and here we want to revisit and recast our original claims. Upon reflection, we think this formulation (which Fourcade and Healy also call attention to) is misleading. If we clarify the reasons why, we can recover the way in which there *is* a kind of classification happening even in scoring models, albeit a classification that is not likely to result in the formation of groups in a sociological sense, per our original argument and consistent with Brubaker's critique.

Let's remember how these two forms of pricing—class-based pricing in insurance and attribute-based pricing in credit—work in practice. (For convenience, we've reproduced the original figures from our paper replicating the insurance pricing table and early credit scorecard in **Figures 1** and **2** below.) To price risk, insurers group

YOUTHFUL OPERATOR
Not eligible for Good Student Credit

AGE			UNMARRIED FEMALE: Pleasure Use or Farm Use	UNMARRIED FEMALE: Drive to Work or Business Use	MARRIED MALE: Pleasure Use or Farm Use	MARRIED MALE: Drive to Work or Business Use
WITHOUT DRIVER TRAINING	17 or Less	Factor Code	1.75 8211— —	2.00 8212— —	1.95 8311— —	2.20 8312— —
	18	Factor Code	1.60 8221— —	1.85 8222— —	1.85 8321— —	2.10 8322— —
	19	Factor Code	1.50 8231— —	1.75 8232— —	1.75 8331— —	2.00 8332— —
	20	Factor Code	1.25 8241— —	1.50 8242— —	1.65 8341— —	1.90 8342— —
WITH DRIVER TRAINING	17 or Less	Factor Code	1.60 8261— —	1.85 8262— —	1.70 8361— —	1.95 8362— —
	18	Factor Code	1.50 8271— —	1.75 8272— —	1.65 8371— —	1.90 8372— —
	19	Factor Code	1.40 8281— —	1.65 8282— —	1.60 8381— —	1.85 8382— —
	20	Factor Code	1.20 8291— —	1.45 8292— —	1.55 8391— —	1.80 8392— —

Figure 1: Class-Based Pricing (Insurance Pricing Table).
Source: Government Accounting Office (1979).

Example of Application Scoring Table						
Years on Job	Less than 6 Months **5**	Six Mos to 1 Yr 6 Mos **14**	1 Yr 7 Mo to 6 Yr 8 Mo. **20**	6 Yrs 9 Mo to 10 Yr 5 Mo. **27**	10 Yrs 6 Mos or More **39**	
Own or Rent	Own or Buying **40**	Rent **19**	All Other **26**			
Banking	Checking Account **22**	Savings Account **17**	Checking and Savings **31**	None **0**		
Major Credit Card	Yes **27**	No **11**				
Occupation	Retired **41**	Professional **36**	Clerical **27**	Sales **18**	Service **12**	All Other **27**
Age of Applicant	18 to 25 **19**	26 to 31 **14**	32 to 34 **22**	35 to 51 **26**	52 to 61 **34**	62 and Over **40**
Worst Credit Reference	Major Derogatory **-15**	Minor Derogatory **-4**	No Record **-2**	One Satisfactory **9**	Two or More Satis-factory **18**	No Investig. **0**

Figure 2: Attribute-Based Pricing (Credit Scorecard).
Source: Lewis (1994).

individuals by identifying characteristics believed to be associated with risk (here, of an automobile accident) and then observe the prior loss history (the actual number of accidents) for this group. On this basis, insurers form an estimate of expected future claims, and they share this cost equally across group members who are assumed to represent equivalent risks to the insurer. Creditors approach the same problem—determining the risk of default on a loan—very differently. Rather than placing each individual into a group with like others, applicants for credit are assigned points on a number of variables, each of which predicts the likelihood of repaying a loan. Tallying up these points results in a score that will be above or below the threshold for receiving credit (set differently by each creditor, as Brubaker notes). Again, the critical idea here is that while an individual applicant for credit might have the same total score as others also applying for credit, they arrive at this score by a different combination of values on variables scored. On this basis, we write in our paper that while “two individuals with the same credit score may share an outcome . . . *they do not share a social experience*” (Krippner and Hirschman 2022, 708; italics added).

The actual mechanics of these two classification systems matter here. When we talk of individuals in insurance markets occupying a class “together,” there is quite literally a cell (i.e., a risk class) in a table that “holds” those individuals (see **Figure 1**). A decision about how to price risk is made for each such cell in the table (even if different cells may be assessed equivalent rates, as Fourcade and Healy correctly observe). Superficially, the credit scoring table looks analogous, but as we note in our paper, each cell in the credit scoring table does not hold a group of individuals, or even a single individual. It holds an *attribute* (a value on a variable used to predict creditworthiness; see **Figure 2**).[20] As these attributes are summed to arrive at a numerical score that determines whether access to credit is granted or denied, the cell in the credit scoring table is *not* strictly analogous to the cell in the insurers' risk pricing table: it does not define the point where the decision (to grant or deny access to credit) is made.

What is important here vis-à-vis Brubaker's critique is that if we are thinking about classification not in terms of what is contained in each cell of the scoring table but in terms of how the decision to grant or deny credit is made by summing *across these cells*, the credit score *does* arguably achieve a kind of classification. That is, any given numerical score will be held by *multiple individuals* who may in this sense constitute a "class" (albeit one with very different characteristics than the class constituted by insurance pricing). Put differently, what is "unique" here is not an individual's outcome on a decision variable (i.e., a score determining access to credit), but the particular *path* by which they arrive at this outcome (or score). This, we believe, is consequential for political mobilization: a shared path is more likely to facilitate political action than a path that is idiosyncratic to a particular person. By the same token, the statement we make in the paper that individuals scored by creditors "may share an outcome, but they do not share a social experience" no longer appears quite correct. We meant to suggest that the outcome (i.e., access to credit) is not attached by the credit score to a shared social position (i.e., as defined by age, gender, marital status, and so on) in the same way that is true for insurance pricing. But the fact of gaining or (especially) being denied access to credit *is* itself a social experience—one that is, in almost all instances, shared with others. And this social experience may in fact constitute a group (e.g., "subprime borrowers") that might be mobilized politically,[21] although we think this is less likely to the extent that this experience is not linked to a fixed location in social space.

We can express this in a way that we think clarifies matters by considering whether the class in question exists independently of the decision technology or is produced by it.[22] Generally speaking, classification in insurance constitutes a group formed by social categories that exist in the world independently of the technology

[20] The image we use in the paper is that of an individual straddling many "subgroups" created by each attribute scored, fragmenting the person and impeding their ability to find community with others (Krippner and Hirschman 2022, 710).

[21] We gave passing attention to the possibility that the score itself could construct a sociologically meaningful group in the conclusion of our paper (see Krippner and Hirschman 2022, 718), but we now think this discussion merits more sustained treatment.

[22] We are grateful to Junchao Tang for help in formulating the points made in this paragraph.

that produces a decision. In credit markets, by contrast, classification constructs a group (defined by the score) that is produced by the decision technology *itself*. To be sure, the creditor's score could, in theory, rely on the same pre-existing social categories that construct the insurance risk class as variables to be scored.[23] But unlike in the insurance case, the attributes used in scoring models are not held constant across individuals, *hence they do not define membership in the group*. Accordingly, we take Brubaker's point that scoring technologies do not involve classification in the sense explored by a now-expansive literature in sociology examining durable forms of social difference (e.g., Tilly 1998; Brubaker 2005, 2016; Massey 2007; Loveman 2014; Mora 2014; Rodriguez-Muñiz 2021; Currah 2022). But this is the case not because each individual gets their own "unique score"—as we wrote in our paper—but because the score forms a group detached from the legible social categories that construct class-based pricing.

Here we might simply adopt Brubaker's (2015) distinction between "categorical" and "gradational" ways of organizing social difference—treating only the former as a "classification"—and call it a day. While we find useful the conceptual distinction between categorical and gradational forms of difference, we aren't quite ready to let go of thinking of the score as a classification. This is because, as we've suggested, the credit score *may* potentially constitute a class (by collapsing a grade into a category)—a "subprime borrower" is not a "nonce-classification" in the sense that Brubaker means. But to emphasize our original point, and in full agreement with Brubaker, neither is the credit score *the same kind of object* as the insurer's risk class. In our view, these are distinct because these two technologies constitute groups in very different ways. To reiterate the key distinction, the insurance pricing table forms a group from social categories that exist independently of the technology that produces a decision (e.g., age, gender, marital status, etc.), whereas the credit score produces a group (if it does) by constructing categories that exist only by virtue of the decision technology itself (e.g., "prime" and "subprime borrower"). Critically, this means that class-based systems will tend to draw on categories that represent hardened forms of social difference. By contrast, attribute-based systems will construct categories *de novo* that will be more tenuous as vehicles for self-identification and consequently social struggle.[24]

[23] "In theory" is important here, as creditors typically do not do this, albeit for historically contingent reasons. Scoring technologies were adopted in credit markets in response to the passage of anti-discrimination laws that expressly prohibited the use of social categories such as gender and marital status routinely used in insurance pricing (see Krippner 2017).

[24] This relates not only to the generally lower social salience and durability of the categories produced by scoring but also reflects the fact that the score is moralized as reflecting one's *own individual choices and behaviors* rather than group memberships (see Fourcade and Healy 2013; 2017).

This brings us to Fourcade and Healy, who make two main points, one which we largely cede, and another which we will attempt to rebut at least in part. The point we cede (as evidenced in our response to Brubaker): there is more variability around the politics generated by attribute-based systems of classification than we allow in our paper. Here, our effort to crystallize the differences between the two classification technologies we examine gave an overly deterministic flavor to our argument. We meant to suggest not that attribute-based decision technologies make collective mobilization impossible, but that they make collective mobilization *less likely* for all the reasons we have just elaborated. Although we want to avoid a deterministic reading of our argument, we also want to underscore that *less likely* is still an important patterning to the politics of the digital age. In this regard, while there is a great deal of variation in how political struggles unfold in a society organized around scoring technologies (which Fourcade and Healy elucidate in their rich discussion), there are also strong tendencies toward depoliticization that we can and should examine.

The second issue that Fourcade and Healy raise falls under their discussion of "duality." Here we have several responses. On its own terms, the idea of duality is very appealing. It captures something profound about the nature of insurance, a long-neglected object of study in sociology that is indeed "dual" in how it deals with individuals and collectives (Ewald 2020). We also appreciate the broader point that attributes and classes can be considered "dual" to each other: attributes can be expressed as classes, and classes as attributes (a point Brubaker also observes in passing). But this observation does not mean that classes and attributes are simply "aspects of the same thing." The act of "conversion" from attribute to class (or from class to attribute) is critical, and it depends in important ways on the nature of the technology used to produce a decision.

We can go a bit further here. Fourcade and Healy make this argument most forcefully by presenting an extended mathematical illustration in which they show very precisely how classes can be expressed as attributes and how attributes can be aggregated into classes. As a purely logical exercise, the demonstration is irrefutable. On another level, however, Fourcade and Healy conflate the mathematics of classes versus attributes with what we might think of as their *phenomenology*. What can be done with a mathematical calculation is very different than what can be done in the world, where conventions, norms, knowledge infrastructures, rules and regulations, and other sociological and legal features shape social practices. Insurance is particularly loaded down by such constraints; as a result, it is perhaps not surprising that although insurers *could* in practice convert classes into attributes, they typically do *not*. Fourcade and Healy write that "if early insurance companies had been able to easily shuffle and swap their categorical configurations, they probably would have.” But as we know from Dan Bouk’s (2015) work on the history of risk-making practices in the United States, insurers at the turn of the twentieth century already had available technologies that allowed them to price individual risk factors (i.e., attributes). However, insurers applied these techniques only in very limited ways because they were at odds with deeply entrenched cultural understandings, professional norms, and regulatory rules that constituted the *risk class* as the basic organizational (and operational) unit in insurance markets (Krippner 2019; 2024).

More surprising, we can observe something similar at play in insurance markets today, notwithstanding new machine learning technologies that have vastly advanced capacities to predict and price individual risks. It is often taken for granted that machine learning and artificial intelligence have transformed pricing in insurance, effectively ushering in a new regime of “data” that has displaced probabilistic (i.e., class-based) methods with fully individualized (i.e., attribute-based) understandings of risk (Ewald 2012). But notwithstanding considerable hype around “personalized” pricing in these markets, a robust literature describes “the revolution that did not happen” in applying machine learning to insurance pricing (Francois and Voldoire 2022; cf., Meyers and Hoyweghen 2018, 2020; McFall 2019; Barry and Charpentier 2020; Cevolini and Esposito 2020, 2022; Jeanningros and McFall 2020; Francois 2025). The reasons for this failed revolution are multi-faceted (see Charpentier and Vamparys 2025), but broadly they reflect basic incompatibilities between the notion of risk pooling (which requires the formation of a group) and individualized pricing reliant on behavioral data. Accordingly, when machine learning is introduced into insurance markets, it tends to be layered on top of traditional actuarial (i.e., class-based) techniques: individuals are first assigned to a class defined using standard demographic variables, and then behavioral data may be applied to make incremental adjustments to this base price. Effectively, to the extent risks are individualized in insurance, it is through the creation of classes that are more finely segmented, not though the overthrow of the class altogether (Barry 2020).

For this reason, we are reluctant to accept Fourcade and Healy’s suggestion that what matters for our argument is not whether data are “statistically represented” in class or attribute form but instead the volume, velocity, and especially behavioral quality of these data. While we certainly agree that the volume, velocity, and behavioral quality of data in modern scoring systems is disabling for political contestation, we do not see this as fully separable from the class or attribute form of the data. This is because whether data are organized around classes or attributes is not a simple matter of how they are “represented”—these are objects in the world that have a tangible reality, not mere representations of that reality (cf., Panagia 2021). In this regard, class-based systems tend strongly toward reliance on a finite number of stable categories, not only for computational reasons but more critically for cultural ones (Krippner 2024).[25] The insurance risk class is derived from traditions of mutualism in which the risk pool constitutes a "community of fate" established between members of a class connected to each other through socially legible characteristics. The attributes scored by creditors create no such community as each individual navigates a singular trajectory through the scoring table. As a result, attribute-based systems are not constrained to select a few characteristics that connect scored individuals in sociologically meaningful ways (although scores may become sociologically meaningful after the fact, as we have noted). Attribute-based systems are, in this regard, more amenable to the "churning" associated with the large volume and high velocity of data used in machine learning applications, with predictable results in terms of the likelihood of collective action.

There is an important caveat to our argument. The features that we attach to class-based systems may be especially marked in insurance, an archaic social institution in which the *risk class* has a special, almost mystical, significance (Krippner 2024). Thus, while insurance pricing has been resistant to transformation through engagement with "big data," the other examples we give in the paper of class-based systems of valuation—marketing and political polling—appear much more permeable to machine learning techniques (see Cheney-Lippold 2011, 2017; Moor and Lury 2018; Brubaker 2023). Even if our argument about class-based pricing applies only to insurance and not more broadly, we still think that our main purpose of illuminating the characteristics of scoring technologies that tend toward disabling political action will have been well served by finding a comparison case that represents the opposed tendencies in their most distilled and potent form.

There is one point on which Brubaker, Fourcade and Healy, and we ourselves converge: the rise of scoring technologies is a deeply ambivalent phenomenon, promising relief from the categorical (class-based) forms of inequality that stratified pre-digital societies, only to reinscribe these inequalities in new, perhaps more pernicious forms (cf., Farrell and Fourcade 2023; Norris Forthcoming). In the face of such challenges, the conceptual clarity provided by Brubaker's and Fourcade and Healy's probing analyses has much to offer the study of politics in the brave new world of the score.

[25] See Gamerdinger (forthcoming) for an account of how cultural differences between actuaries and data scientists have operated to slow the adoption of machine learning within insurance organizations.

Acknowledgements

We thank Junchao Tang for helpful comments on an earlier version of this essay.

Competing Interests

The authors have no competing interests to declare.

Sources Cited

Barry, Laurence. 2020. “Insurance, Big Data, and Changing Conceptions of Fairness.” *European Journal of Sociology* 61(2): 159–84.

Barry, Laurence, and Arthur Charpentier. 2020. “Personalization as a Promise: Can Big Data Change the Practice of Insurance?” *Big Data & Society* 7(1): 1–12.

Bouk, Dan. 2015. *How Our Days Became Numbered: Risk and the Rise of the Statistical Individual*. University of Chicago Press.

Breiger, Ronald L. 1974. “The Duality of Persons and Groups.” *Social Forces* 53: 181–90.

Brubaker, Rogers. 2005. “Ethnicity without Groups.” In *Remaking Modernity*, edited by Julia Adams, Elisabeth S. Clemens, and Ann Shola Orloff. Duke University Press.

Brubaker, Rogers. 2015. *Grounds for Difference*. Harvard University Press.

Brubaker, Rogers. 2016. *Trans: Gender and Race in an Age of Unsettled Identities*. Princeton University Press.

Brubaker, Rogers. 2023. *Hyperconnectivity and Its Discontents*. Polity Press.

Cevolini, Alberto, and Elena Esposito. 2020. “From Pool to Profile: Social Consequences of Algorithmic Prediction in Insurance.” *Big Data & Society* 7(2): 1–11.

Cevolini, Alberto, and Elena Esposito. 2022. “From Actuarial to Behavioural Valuation: The Impact of Telematics on Motor Insurance.” *Valuation Studies* 9(1): 109–139.

Charpentier, Arthur, and Xavier Vamparys. 2025. “Artificial Intelligence and Personalization of Insurance: Failure or Delayed Ignition?” *Big Data & Society* 12(1): 1–13.

Cheney-Lippold, John. 2011. “A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control.” *Theory, Culture, and Society* 28(6): 164–181.

Cheney-Lippold, John. 2017. *We Are Data: Algorithms and the Making of Our Digital Selves*. NYU Press.

Currah, Paisley. 2022. *Sex Is as Sex Does: Governing Transgender Identity*. NYU Press.

Ewald, Francois. 2012. “Assurance, Prévention, Prédiction: Dans L’univers du Big Data.” Institut Montparnasse. https://www.institutmontparnasse.eu/wp-content/uploads/2012/11/Collection_recherches_n_4.pdf.

Ewald, Francois. 2020. *The Birth of Solidarity: The History of the French Welfare State*. Duke University Press.

Farrell, Henry, and Marion Fourcade. 2023. “The Moral Economy of High-Tech Modernism.” *Daedalus* 152(1): 225–235.

Fourcade, Marion. 2016. “Ordinalization.” *Sociological Theory* 34(3): 175–195.

Fourcade, Marion, and Kieran Healy. 2013. “Classification Situations: Life Chances in the Neoliberal Era.” *Accounting, Organizations and Society* 38(8): 559–72.

Fourcade, Marion, and Kieran Healy. 2017. “Seeing Like a Market.” *Socio-Economic Review* 15(1): 9–29.

Fourcade, Marion, and Kieran Healy. 2024. *The Ordinal Society*. Harvard University Press.

Francois, Pierre. 2025. “The Old Regime (of Mutualisation) and the Revolution (of Big Data).” *The British Journal of Sociology* Special Issue: Insurance and Social Theory: 1–8.

Francois, Pierre, and Theo Voldoire. 2022. “The Revolution that Did Not Happen: Telematics and Car Insurance in the 2010s.” *Big Data & Society* 9(2): 1–12.

Gamerdinger, Alexander. Forthcoming. “Moral Authority Over Risk Classifications: How Data Professionals Shape the Uneven Algorithmization of Life Insurance.” *Socio-Economic Review*.

Government Accounting Office. 1979. *Issues and Needed Improvements in State Regulation of the Insurance Business: Report to the Congress*. Washington, D.C.: Government Printing Office.

Jeanningros, Hugo, and Liz McFall. 2020. “The Value of Sharing: Branding and Behaviour in a Life and Health Insurance Company.” *Big Data & Society* 7(2): 1–15.

Krippner, Greta R. 2017. “Democracy of Credit: Ownership and the Politics of Credit Access in Late Twentieth-Century America.” *American Journal of Sociology* 123(1): 1–47.

Krippner, Greta R. 2019. “Bearing Risks, Being Watched.” *Public Books*, February 26. https://www.publicbooks.org/bearing-risks-and-being-watched/.

Krippner, Greta R. 2024. “Gendered Market Devices: The Persistence of Gender Discrimination in Insurance Markets.” *American Journal of Sociology* 130(3): 595–643.

Krippner, Greta R., and Daniel Hirschman. 2022. “The Person of the Category: The Pricing of Risk and the Politics of Classification in Insurance and Credit.” *Theory and Society* 51(5): 685–727.

Lewis, E. M. 1994. *An Introduction to Credit Scoring*. Athena Press.

Loveman, Mara. 2014. *National Colors: Racial Classification and the State in Latin America*. Oxford University Press.

Massey, Douglas S. 2007. *Categorically Unequal: The American Stratification System*. Russell Sage Foundation.

McFall, Liz. 2019. “Personalizing Solidarity? The Role of Self-Tracking in Health Insurance Pricing.” *Economy and Society* 48(1): 52–76.

Meyers, Gert, and Ine Van Hoyweghen. 2018. “Enacting Actuarial Fairness in Insurance: From Fair Discrimination to Behaviour-based Fairness.” *Science as Culture* 27(4): 413–438.

Meyers, Gert, and Ine Van Hoyweghen. 2020. “Happy Failures: Experimentation with Behaviour-Based Personalisation in Car Insurance.” *Big Data & Society* 7(1): 1–14.

Moor, Liz, and Celia Lury. 2018. “Price and the Person: Markets, Discrimination, and Personhood.” *Journal of Cultural Economy* 11(6): 501–513.

Mora, Cristina G. 2014. *Making Hispanics: How Activists, Bureaucrats, and Media Constructed a New American*. University of Chicago Press.

Norris, Davon. Forthcoming. “Epistemological Sleights or How Racial Capitalism Gets Away with Being Racial.” In *Political Economy, Rebooted*, edited by Marion Fourcade, Greta Krippner, and Sarah Quinn. Duke University Press.

Panagia, Davide. 2021. “On the Possibilities of a Political Theory of Algorithms.” *Political Theory* 49(1): 109–33.

Poon, Martha. 2012. *What Lenders See: A Pre-History of the FICO® Credit Scoring System*. Ph.D. Thesis, University of California, San Diego.

Rodríguez-Muñiz, Michael. 2021. *Figures of the Future: Latino Civil Rights and the Politics of Demographic Change*. Princeton University Press.

Tilly, Charles. 1998. *Durable Inequality*. University of California Press.