

# What To the Educational Researcher Is A Stylized Fact?

Derek Gottlieb and Jack Schneider

*University of Northern Colorado and University of Massachusetts Amherst*

## INTRODUCTION

This talk anatomizes the insidious role that stylized facts play in the public politics and policy apparatus of education reform. A stylized fact, as the sociologist Dan Hirschman describes it, is most fundamentally a certain framing of “an empirical regularity in need of explanation.” As such, stylized facts are particularly useful to a political culture committed to sharp distinctions between fact and value, between theory and practice, and between description and advocacy. Stylized facts straddle each of these dichotomies in such a way that they can perform functions in one realm—fortifying normative commitments, warranting certain practical courses of action, and advancing large-scale ideological frameworks—while plausibly denying that they are anything more than merely factual descriptions of the world suitable for theorizing.

The central place that stylized facts have come to occupy in educational research and policy is bad for several reasons. It short-circuits the deliberation crucial to democratic life by obscuring the nature of normative disagreements and concomitantly shoring up a technocratic idealization of the social and political world. It encourages material and affective overinvestment in small-scale changes to policy structures or practical techniques by furnishing unwarranted confidence in predicted large-scale outcomes. And, putting both of the previous points together, stylized facts drive “reform churn,” the successive parade of policy interventions that traffic in normative appeals to equality and justice while simultaneously undermining institutional stability through abrupt technocratic rewritings of regulatory infrastructure.

## WHAT STYLIZED FACTS ARE

The sociologist Dan Hirschman has done the most comprehensive analysis of stylized facts to date, and our exploration of the matter seeks to expand upon his account. In particular, we wish to detail more fully the nor-

mative dimensions and applications of stylized facts in the educational policy and research ecosystem.

Hirschman credits the economist Nicholas Kaldor with coining the term in 1961, quoting Kaldor as saying that “the theorist, in my view, should be free to start off with a ‘stylized’ view of the facts—i.e. concentrate on broad tendencies, ignoring individual detail. . . [and] construct a hypothesis that could account for these ‘stylized facts,’ without necessarily committing himself on the historical accuracy, or sufficiency, of the facts or the tendencies thus summarized.”<sup>1</sup> Kaldor’s version of a stylized fact frees theorists from bogging down in accounting for variation across particular cases that may reflect nothing more consequential than measurement error so that they can get on with their theory-building work.

Hirschman tacks on another normative benefit of stylization to Kaldor’s description, as well. He notes that stylized facts solve a finite-attention problem for social scientists: they amount to normative claims about which tendencies and social kinds are most deserving of our attention and study.<sup>2</sup> Stylized facts are not only internally useful to academic disciplines as rough-and-ready grist for theory-construction, but they are also externally useful to social scientists as a means of promoting a particular empirical regularity as a matter of public concern or of translating vague public concerns into the terms of an empirical regularity operationalized for scientific exploration.

Following this expansion on Kaldor, Hirschman identifies four key features of stylized facts:

1. “They presume or create an analytical ontology,” which is to say that they presume or create “the stable existence of social kinds worth theorizing about.”<sup>3</sup> Existence, stability, and the notion of “worth” are all key elements here.
2. These social kinds “are defined in terms a non-specialist can understand,” which makes social-science research susceptible to—or available for—productive misunderstandings. Specifically, social scientists must “strategically bracket” the ordinary complexities and historical variation

internal to social kinds in order to operationalize definitions amenable to scientific inquiry (race, student learning, etc.). But because the *names* of these simplified social kinds circulate widely and variously in everyday discourse, the conflation of technical and ordinary meanings is always possible, if not actually invited.<sup>4</sup>

3. To quote Hirschman directly, “stylized facts are typically understood as claims of non-robust dependence.”<sup>5</sup> They are empirical regularities, “simple associations,” whose regularity signals some sort of dependence relation. The nature, direction, strength, and/or underlying mechanisms of that relation becomes the target of further social-science research. This is one area of Hirschman’s analysis that we wish to augment: it has lately become typical for stylized facts in educational research and policy to take the form of *robust*-dependence claims, which only amplifies their social and political power.
4. Stylized facts are always normative claims in at least two ways: they are claims about which regions of social life are most worthy of our finite attention, and they are claims about which theory-laden description of a given regularity is the best or most appropriate way of characterizing a state of affairs. Here, too, we will offer an augmentation of Hirschman’s description by noting a third aspect of stylized facts’ normativity: they make claims about which elements of social or institutional structures are realistically alterable and therefore appropriate targets for policy intervention.

We can summarize the major contours of this overview by saying that the *stylization* of facts, in both Kaldor and in Hirschman, involves the temporary suspension of real-world messiness in order to articulate a larger and more internally consistent truth for public consumption and expert study. The simplification of complex historical processes into broad tendencies relating social categories of public importance simultaneously suits such stylized facts for scientific and political projects. The capacity of stylization to blur the distinction between disinterested scientific fact and partisan political value lends public prestige—and the material benefits that flow from it—to social-scientific find-

ings, as this capacity also furnishes normative political agendas with (apparently) factual warrant. This distinction-blurring capacity comes with enormous power.

It also comes with enormous potential for witting or unwitting abuse, for sowing public confusion, for undermining democratic habits, and for misdirecting material and affective resources on the basis of overstated or undertheorized promises. Understanding the harms that can result, and that have resulted, from enshrining stylized facts as the primary currency in the research-and-policy ecosystem requires zeroing in on the way stylized facts make themselves “capable of being (productively) misunderstood,” in Hirschman’s wonderful phrase.<sup>6</sup>

Let us take a paradigm case from our own domain of educational research and policymaking to illustrate the issues involved.

#### THE CASE OF RACE TO THE TOP

In his 2012 State of the Union Address, President Obama said the following: “We know a good teacher can increase the lifetime income of a classroom by over \$250,000.”<sup>7</sup> He takes the truth of this assertion from the abstract of Raj Chetty’s groundbreaking 2011 work on the value-added impacts of teachers, in which Chetty wrote, verbatim: “Replacing a teacher whose VA [value-added] is in the bottom 5% with an average teacher would increase the present value of students’ lifetime income by more than \$250,000 for the average classroom in our sample.”<sup>8</sup>

In both of its forms, whether Obama’s or Chetty’s, this is a stylized fact *par excellence*. It meets all of Hirschman’s criteria. It is an empirical regularity in need of explanation, which suits it for uptake in the social-science research apparatus: we need to know why this is the case, we need to know what good teachers are doing that mediocre teachers are not, and so on. Voila—a research agenda.

It also breaks the social world down into particular stable kinds and suggests that certain ones amongst them are most worthy of attention. “Good teachers” are constructed as a unitary category whose significant variation across historical and geographic context is ignored for the sake of expediency,

which adds the impression of stability to the obvious implication of existence. The assertion of the fact itself signals its public importance. Good teachers and lifetime earnings are things that we do (and ought to) care about, and the additional information that they are associated by in some regular way only gives us more reason to attend to them. Good teachers are not only intrinsically worthy of our attention but instrumentally worthy, too.

The intrinsic and instrumental value of good teaching reflects a third characteristic in Hirschman's definition: the tendency of a stylized fact to use misleadingly ordinary terms. "Good teachers" and "lifetime income" are immediately comprehensible publicly in ways that allow or invite the overlooking of technical caveats and definitional qualifications. In fact, Obama's *good* teachers already depart from the "high-VA" teachers that Chetty explicitly names in his abstract. Later in that same State of the Union address, Obama goes as far as saying that "every person in this chamber" can remember such a teacher in their own lives, thus completing the erasure of any distinction between Chetty's technical definition of quality—a tendency to cause students to perform better on state tests of math and reading in grades three through eight than prior performance and demographic factors would predict—with a much broader sense of good teaching rooted in the "trajectory-changing" effects of the pedagogical relationship. This is an instance of what Hirschman calls stylized facts' capacity to "travel" across discursive boundaries, migrating back and forth between the narrowly described confines of social-science research and the maximally broad arena of the public interest. This traveling capacity produces a general overestimation of the policy-relevance of social-scientific findings, as witnessed in Obama's own conceptual slippage.

Our example of a stylized fact makes normative claims in both of the ways Hirschman identifies, as well. It directs our attention to particular regions of social life as most worthy of public consideration, as we have noted, and it further makes the implicit assertion that it is the most appropriate characterization of the regularity in question. An economic metric like lifetime income insists that it is, rhetorically and scientifically, in the final analysis, the ultimate product of an educational process. It therefore also suggests that education

ultimately or fundamentally serves an economic function. Beyond this narrowing or simplification of the terms in question—the significance of lifetime income and of education—this stylized fact also implicitly proposes itself as superior to alternative characterizations of the same regularity. That is, the same dataset that Chetty used, parsed with same methods of inference, would be captured equally well in the following stylization: “Our macroeconomic institutions and policies reliably distribute lifetime income across the population according to individual levels of performance on tests of math and reading taken during grades three to eight, which can be influenced to varying degrees by the particular teachers to which students are assigned.” Stylizing an empirical regularity one way rather than another is a normative act of judgment that focuses public and political attention on certain aspects of the social ecosystem rather than others. This act of judgment, as Hirschman describes it, is harmful because it excludes potential alternatives before bringing its case to the public, so to speak.

Recognizing the normative work done by our paradigmatic instance of a stylized fact leads us to revisit and update Hirschman’s final characteristic of stylized facts, namely, the non-robust dependence claim. Strictly speaking, as we will show, Chetty’s assertion about the \$250,000 in excess income is a claim of *robust* dependence, produced by appropriately rigorous inferential methods. Nevertheless, Obama’s articulation of it (“we know a good teacher can increase the lifetime income, etc.”) goes well beyond the warrant that Chetty’s research provides, and not merely because of the way it implies a specifically causal relation.

The prototypical claims of non-robust dependence that Hirschman describes do not invite this kind of policy-relevant misunderstanding. Take, for instance, Max Weber’s observation that Protestant regions seem to take to capitalism with unusual alacrity. While Weber’s observation does similar normative work—implicitly centering religion and economic structure as matters of sociological attention—and while there are surely other ways of characterizing the relation, the observation itself is primarily useful as an impetus to theorizing. Even if there was a political interest in promoting capitalism, the simple association between Protestantism and capitalism provides no reason to believe that increasing Protestantism will increase capitalism without more theorizing

and empirical research.

Chetty's claim is different, however. He claims to have discovered, minimally, that it would have been possible, under certain conditions, to change lifetime earnings by swapping teachers of different levels of effectiveness. It is a much smaller step, and a much more tempting one, for policymakers interested in increasing lifetime earnings to simply project the continuity of this robust dependency into the future, perhaps with the prudent caveat that the size of the impact might change with contextual variation. The discovery of a relation of robust dependence under specific historical conditions is especially susceptible to unwarranted uptake in political efforts to influence future circumstances in particular ways. Our example shows that if ambiguously ordinary terminology allows social-scientific findings to travel across discursive boundaries, then ambiguously robust dependence claims allows social-scientific facts to travel across temporal boundaries, as well, in a way that Hirschman lets alone. Let us draw this out a little further.

The Chetty finding, and Obama's citation of it, is produced by logistic regression procedures explicitly designed to control for potential mediators or confounders. This does, in fact, warrant a stringently-qualified claim of *robust* dependence. Specifically, in this instance, Chetty establishes the relation between good teaching and lifetime earnings by comparing scores on tests taken between 1989 and 2009 against class rosters, tax records, and so on, from the same period. The robustness of this dependence relation only applies, and can only be inferred or suggested, within the particular macroeconomic and geopolitical conditions of the US situation between the end of the Cold War and the Great Recession, roughly. This is why Chetty takes care, in his own stylization of the fact, to say that *if* a particularly poor teacher *had been* replaced with a merely average teacher, the lifetime income of a classroom *in their sample* would have increased by \$250,000. That, strictly speaking, is the relation of robust dependence warranted by Chetty's research, articulated as a stylized fact. Its robustness depends upon its backward-facing perspective. So when this fact is articulated in a State of the Union address, it is *deploying* a real but massively-constrained claim of robust dependence in the direction of the future, where no such robust

dependence can simply be assumed to obtain.

That is what makes the connection between good teaching and lifetime earnings a claim about what we call *ambiguously* robust dependence. We know that it *is* robust in highly circumscribed contexts that have no political utility since we cannot change the past. We do *not* know that it is robust in the future contexts of interest to policymakers and the public. But we also do not strictly know, pending further research into the relation itself, that it will *not* be equally robust in the future. Chetty's stylized fact, as taken up by President Obama, embodies the *suggestive* finding, whose expertly produced warrants in tightly-defined circumstances epistemologically underwrite the shape and target of policy interventions aimed at certain normative ends. When Obama turns Chetty's highly conditional articulation of what *would* have happened to the lifetime earnings of *their sample* if one teacher rather than another stood at the front of the class into a simple truism that "we know a good teacher can increase the lifetime income of a classroom by \$250,000," he erases the *pastness* of the dependency and, thereby, inaccurately characterizes Chetty's finding in a way that cues up a suite of policy options for achieving a normative end. That is what a productive misunderstanding looks like.

### PRODUCTIVE MISUNDERSTANDINGS

We have already explained the "misunderstanding" half of the term by documenting the capacity stylized facts' warrants and terminology have for "traveling" across space and time, allowing various social actors to misapprehend stylized facts as unproblematically applicable to circumstances beyond the explicit reach of a given social-scientific finding. These misunderstandings are also *productive*. Specifically, stylized facts are selective in highlighting and downplaying their normative elements in ways that tacitly privilege certain institutional arrangements. By concealing some of the normative work that they do, stylization allows a fact to seem *more* ideologically inert than it is. Stylization prevents a full public consideration of the normative elements involved. This aprioristically limits the shape and the range of reasonable political problems and their solutions. So stylized facts are productive in the sense of setting up or shoring up a particular kind of public reason that structures the distribution and



the targeting of resources in particular ways. Let us take the example of good teachers and lifetime incomes again, though more briefly this time.

Good teachers increase lifetime earnings. This fact selectively highlights the normative claim that increasing lifetime earnings is good, which is a claim that can accommodate a range of moral frameworks, from a basic individual-utility perspective to a capability's perspective, and so on. Our stylized fact also draws more implicitly on education's historical place in this country in particular as a technology of both freedom and domination, characterized and carried out by the differential provision of funding and resources, including good teachers. Higher lifetime income is generally good for everyone, and the wider availability of good teachers is highly responsive to a legacy of institutional inequality. But this fact also selectively *downplays* the normative judgment to make teacher quality the focus of analysis, rather than, as we have suggested, the macroeconomic arrangements that distribute incomes according to educational credentials of various kinds. The production of a stylized fact relating teacher quality to lifetime incomes performs the normative operation of targeting the teaching profession for policy intervention. It gives policymakers something specific to do.

The unacknowledged normative effects of stylization are more far-reaching, too. Our paradigmatic stylized fact posits the vague metric of lifetime incomes as a stand-in for a more comprehensive account of what we value in education and why. It creates the appearance of broad-based moral consensus without having to hash out the particular reasons that different people may have for valuing higher incomes or for looking to education as means of generating them. Perhaps more importantly, this moral consensus appears without having to consider what different people may or may not be willing to entertain as tradeoffs, either for higher incomes in general or for an education system dedicated to the production of higher incomes in particular. The use of a minimally objectionable metric for a minimally objectionable good performs the normative and ontological tasks of simplifying political goals and reducing the salience of political pluralism. It sets the parameters for *reasonable* research and policy agendas in general.

And lastly, the systematic privileging of certain research methods functionally reinforces the normative commitments we have just discussed. As we have said, Chetty's regression analysis enables the inference of a robust dependence relation between teacher quality and lifetime income, as long as that inference is restricted to the general macroeconomic context in which the empirical regularity was observed. As we have also said, the *way* that Obama cites Chetty's work in the State of the Union address transforms an appropriately qualified socioeconomic relation into a promissory article of faith, a specific wager that focusing policy attention on the teaching profession will yield certain material effects that are normatively desirable. Because Obama's promissory rhetoric—and the specific, normatively-grounded assurances of public policy more generally—is founded on the robustness of a relation discovered within highly-particular macroeconomic conditions, the fulfillment of policy promises paradoxically commits us to the indefinite maintenance of those very macroeconomic conditions.

The misunderstandings to which stylized facts lend themselves are productive, then, in these particular ways. The ambiguous nature and quality of a stylized fact's evidentiary warrant augers against "big structural change": acting to alter macroeconomic conditions will positively undermine the predictive knowledge supposedly represented in a given stylized fact. This substantially narrows the range and ambition of "reasonable" policy options for the public to consider. Misunderstanding the full extent of normative work intrinsic to the stylization of facts also produces an overstated sense of overlapping consensus, to coin a phrase, which allows policymakers to simply bypass deliberative processes in the run-up to policy design and implementation. This is generally out of alignment with democratic forms of sociality, mainly because it discourages us from engaging in the public work of mutual accommodation, in which we might understand what different members of our communities have at stake in decisions about whether and by what means to raise lifetime incomes in general.

In sum, stylized facts provide reasons for policymakers and the public to believe that we can meet our collective normative and material needs through minimally invasive tweaks on the margins of our existing institutional

configurations. But our reasonable belief, here, is rooted in the way that stylization works to obscure the extent of our value pluralism, to limit the range of plausible social or regulatory reforms, and to overstate the predictability of future outcomes. Stylized facts are the fuel on which technocratic governance structures run, and they operate to disincentivize the cultivation of democratic habits and practices.

But much more than that, the institutional reliance on policy-relevant stylized facts, on evidence-based policy interventions, produces the destabilizing dynamics of reform churn.

## CONCLUSION

### ESRA, DOUBLING DOWN, AND REFORM CHURN

The lesser-known federal education bill passed in 2002 was the Education Sciences Reform Act, which created IES in order to direct federal resources to policy-relevant educational research. Speaking at AERA the following spring, the first head of IES, Russ Whitehurst, explained what was wrong with education research, how to improve it, and how IES would help. Whitehurst basically told his audience that educational research over the past few decades had become too theoretical, too accommodating of methodological laxity, too divorced from the needs of school personnel, and too ideological. Whitehurst told the story of being flummoxed by a superintendent who asked him what the best math curriculum for fourth graders was. There was no existing research that comprehensively answered the question, he replied, offering his opinion instead. The superintendent said, “I have enough opinions.” Whitehurst then proceeded to lay out IES’s funding criteria, designed to incentivize the methodologically-rigorous production of ideologically-neutral facts focused on specific areas of concern to school system leaders. These privileged areas of concern included disparities in educational funding, the racial achievement gap, and effective instructional techniques. The hierarchy of methodological rigor, unsurprisingly, featured the randomized controlled trial at the top, with quasi-experimental and correlational methods just below, and qualitative methodologies at the bottom. In 2015, the Every Student Succeeds Act inscribed this evidentiary hierarchy

into its regulation of states and local policy-adoption processes.

It is not an overstatement or a mischaracterization to say that IES was founded to promote the production of stylized facts for application in educational policy aiming at particular normative ends. Chetty's study, which we have discussed throughout, epitomizes the work IES has wanted to encourage: it focuses on a privileged area of inquiry, it uses rigorous quantitative inference methods, and it produces an epistemologically-sound finding of a specific relation of robust-dependence with immediate policy implications to which the president himself gave voice. But, to gloss a complex, decade-long history, the effects of implementing teacher-targeting policies based on Chetty's discoveries have been underwhelming in terms of their intended effects and overwhelming in their unintended consequences.

Why? Most basically, it is because of what a stylized fact *is*. A stylized fact is attractive to policymakers and the public because it seems to offer future assurances pertaining to ordinary features of common experience. But its future assurances are subject to rigorous qualification that stylization conceals, and its operationalized definitions of everyday terms necessarily fail to map onto the range of popular usages. Designing a policy around the terms and definitions of a stylized fact and expecting that policy to produce the promised outcomes beyond the context in which the stylized fact was established will *necessarily* fail to achieve the expected results either in specific outcome measures or in the vaguer terms of political accolades.

Stylized facts have not, so far, taken their fair share of the blame for policy disasters. It is much more common to think that we were wrong to focus on this or that element of the education system specifically, or to blame local practitioners for not implementing policy intentions with sufficient fidelity. This, too, is part of the productive misunderstandings generated by the ambiguity of stylized facts themselves. The traveling capacity inherent to stylized facts—their seeming applicability to wider conceptual and temporal contexts—prevents us from recognizing stylized facts themselves as playing a causal role in policy shortfalls and unintended consequences. So the market value of stylized facts in our research and policy ecosystem remains stable. And we remain nearly as

far from realizing our normative commitments as ever. It is this combination of factors—the normative impetus to live up to our political ideals and a set of preferentially-funded research methodologies—that creates the churn cycle, proceeding from research finding to policy design to roll-out and implementation to failure, and on to the next research finding.

Stylized facts constitutively require the erasure of pluralism, whether in terminological understandings or normative frameworks or matters of public concern. And yet our political ideals all take human pluralism as a first fact: however we might understand terms like justice or freedom or equality, we understand them as applying in a pluralistic context. Stylized facts seem to offer us a shortcut through the slow and messy work of democratic decision-making to the outcome on which we all already agree. But the shortcut is achieved by deceptive means, smuggling in certain normative commitments and aprioristically excluding certain political options from consideration. Far from streamlining a process of discovering and selecting the means for achieving a universally agreed-upon end, relying on stylized facts to direct policy design eliminates that process entirely, thus denying the value of *coming* to agreement on matters of public value and *deciding* on the proper course of action from a range of possibilities. Valuing such decision-making practices purely in terms of the particulars of the eventual decisions themselves is how the technocrat mistakes the existential difficulty of living with others for an intellectual puzzle.

---

## REFERENCES

- 1 Nicholas Kaldor, *Capital Accumulation and Economic Growth* (Macmillan, 1961), 178, [https://doi.org/10.1007/978-1-349-08452-4\\_10](https://doi.org/10.1007/978-1-349-08452-4_10); Daniel Hirschman, “Stylized Facts in the Social Sciences,” *Sociological Science*, 3 (2016): 606. <https://doi.org/10.15195/v3.a26>
- 2 Hirschman, “Stylized Facts in the Social Sciences,” 606.
- 3 Hirschman, “Stylized Facts in the Social Sciences,” 607.
- 4 Hirschman, “Stylized Facts in the Social Sciences,” 607.
- 5 Hirschman, “Stylized Facts in the Social Sciences,” 607.

6 Hirschman, “Stylized Facts in the Social Sciences,” 605.

7 Barack Obama, “Remarks by the President in State of the Union Address,” National Archives and Records Administration, accessed November 1, 2023, <https://obamawhitehouse.archives.gov/the-press-office/2012/01/24/remarks-president-state-union-address>.

8 Raj Chetty, John N. Friedman, and Jonah E. Rockoff, “Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood,” *American Economic Review* 104, no. 9 (2014): 2633–79. <https://doi.org/10.1257/aer.104.9.2633>