Errors in school funding pieces that demand transparent correction

1) International comparisons of school funding:

In his initial piece, Mehlhorn wrote the following:

> Finland's per-pupil spending of $10,905 in 2011 is lower than the $15,345 spent by the United States; South Korea's $8,382 per-pupil is 83 percent lower. Based on comparisons with those two countries alone, it becomes clear that money isn't the main problem in American public education.

As I pointed out in my response, the statistics Mehlhorn cites here do not support his point about *public* education:

> The problem, however, is that the numbers in Mehlhorn's piece are cherry-picked; they don't actually speak to his argument about *public K-12* education spending. As the OECD report notes, the figures Mehlhorn cites include public *and private* spending on primary, secondary, *and tertiary education – that is, college* – including but not limited to spending on transportation, meals, school health services, college dormitories, and "private spending on books and other school materials or private tutoring."
>
> In general, the OECD data shouldn't be used for cross-country comparisons; it doesn't count spending the same way in each country and likely makes US spending appear larger relative to spending in other countries than it actually is. To the extent that the data can be illustrative, however, the appropriate approach would exclude college costs and private spending and focus on K-12 public school spending as a share of the economy (as opposed to using raw numbers; spending as a share of GDP provides a better indication of how much a country spends relative to what it can afford). Doing so (see Table B4.1 here) indicates that public spending on primary and secondary education in the United States, relative to GDP, is lower than spending as a share of the economy in Finland, the same as such spending in Korea, and slightly below the OECD average. Again, the data is flawed, but it likely provides a high-end estimate of United States education spending relative to such spending elsewhere.

Mehlhorn never corrects his original piece by noting that his numbers were misleading. He also does not admit in his subsequent posts that he has made a mistake; instead, he says only that I "did not like the source [he] used" and argues that I "bur[y] the international comparisons behind a number of tangents."

Though he at least admits that I am "right that we should try to compare apples to apples," he then proceeds to cite a stat from McKinsey which appears to rely on earlier data produced by the *same source he cited before*. What's more, the spending data for Singapore that he uses is, as McKinsey notes, "relative to GDP per capita," a metric Mehlhorn argues "has uses, but is not relevant to whether schools have enough resources." The NCES link he cites also notes that the data comes "from both public and private sources" and makes the same point I made: that the United States spends less on education relative to the size of its economy than is typical.

Here's Mehlhorn's first response:

> **How much does the United States spend on education?** The starting point, which Ben buries behind a number of tangents, is that America spends more than any society in history. All credible sources tell roughly the same story. Ben did not like the source I used in my first column, so let's use the National Center for Education Statistics. It reports a $621 billion total national investment in public elementary and secondary schools for 2011-2012.
>
> **We spend more than other countries spend, yet get weaker results**. The international story has been repeatedly verified: Compared to other nations, the United States spends more on each student, and the students get less. Ben is right that we should try to compare apples to apples, but the best efforts to do so repeat this conclusion. An older, thorough study by McKinsey & Company in 2007, noted that Singapore achieves top performance while spending less per pupil than 27 of 30 OECD countries. More recently, NCES says we spend $12,401 per pupil, about 35 percent more than the per-pupil average for the industrialized world. In case after case, and in study after study, the best school systems do more with less than America and its public education systems.
>
> Ben's responds that education spending should be measured as a share of Gross Domestic Product, rather than as an absolute number. In this, Ben forgets what we are discussing: whether schools in America have enough money to succeed. His preferred metric – education spending divided by GDP – has uses, but is not relevant to whether schools have enough resources.

And here's his updated response:

Like other reform skeptics, Ben understandably buries the international comparisons behind a number of tangents. The reality is that America spends more on its schools than any society today, or even any society in human history. Yet, we do not achieve commensurate results; on the contrary, educational outcomes in the USA are unfavorable in comparison with high-test, high-stakes countries (Korea), with rich and low-test countries (Finland), and even with poor countries struggling to grow (Poland).

As I wrote in my first update, I'm happy to debate whether *public* spending as a share of GDP per pupil (on which the US is below average) or absolute *public* spending numbers per pupil (on which the US is above average, though Mehlhorn's assertion that it is more "than any society today, or even any society in human history" is false) is a more appropriate comparison. I'm also happy to explain why the logic he uses to link absolute spending numbers and student outcomes is faulty. In fact, I did both in my update:

> One fair point Mehlhorn does make is that inflation-adjusted spending levels have value. I used spending as a share of GDP above to note that the US spends less on education relative to what we can afford than many other countries and that our education spending relative to what we can afford hasn't changed much over time. Those facts in and of themselves don't necessarily mean that our spending levels are insufficient; they just show that our investment in education is consistent with historical and international norms. But while it's fine for Mehlhorn to note that per-pupil spending in the US is up significantly in real terms since the 1970s, *that also doesn't necessarily tell us anything about whether spending levels are sufficient*. We may have been spending way too little in the 1970s, and we still may be spending way too little now.
>
> In any case, Mehlhorn's note that education spending has increased more than test scores doesn't say anything, by itself, about the efficacy of that spending. Student test scores are influenced more by outside-of-school factors than by school-based factors and *it's impossible to know how effective an intervention was without knowing what would have happened in the absence of the intervention*. Maybe test scores would have *fallen* if spending had remained flat. We don't know. What we do know is that studies that attempt to identify a counterfactual, like Jackson et al.'s, indicate that increased school funding makes an important difference.

The thing is, it's hard to have that debate when your debate partner keeps misleading readers with incorrect information.

2) *Historical comparisons of school funding:*

In my original piece, I wrote:

> Mehlhorn's article also paints an incomplete picture of historical levels of education funding in the United States. The fact that K-12 spending has risen in inflation-adjusted dollar value terms over the past 45 years doesn't tell us anything about whether school spending levels are sufficient, and real spending on practically everything has increased in dollar terms since the 1970s; in fact, real spending *should* increase as our economy grows. A more appropriate (though still imperfect; one flaw is

Mehlhorn then wrote the following response in his "rebuttal:"

> **We spend a lot more than we used to, without commensurate results:** America's schools today spend about 2.5 times per pupil what they spent in 1970, notwithstanding a small per-pupil dip since 2008. Ben acknowledges "the fact that K-12 spending has risen in inflation-adjusted dollar value terms over the past 45 years," but then waves that away by saying that "real spending on practically everything has increased in dollar terms since the 1970." That statement is jarringly untrue.
>
> Over that time period, per-unit prices have plummeted in many areas, including appliances, telecommunications, electronics, computers, televisions, and audio-visual devices. Some sectors have taken advantage: for instance, U.S. military spending has increased only 10 percent since 1970, while dramatically improving its comparative and absolute effectiveness. True, declining costs in some sectors have been offset by price increases in other areas, but this overall mix is called "inflation." By using "inflation-adjusted" dollars, we account for the interplay of cost increases and cost declines. If we ignored inflation, the increase in dollars would be 14 times rather than merely 2.5 times.

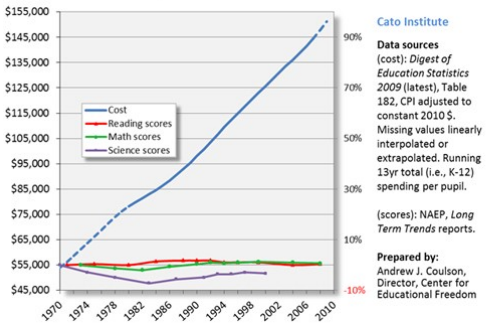
Rather than being "jarringly untrue," however, my original statement is correct – it is Mehlhorn's arguments which are factually wrong, as I explained in my first update:

> 2) David Dayen recently wrote an excellent piece about why citations of raw numbers for government spending – of the type that appear in Mehlhorn's piece – are misleading. I highly recommend it. Mehlhorn is also mistaken about historical trends in real (inflation-adjusted) spending outside of education; as a quick look at the data for some of the categories he mentions (like certain technologies or defense) confirms, spending on (which is different than prices of things in) these categories has also grown over time (though by different amounts than education spending and not on a per capita basis for defense, which it would have been fine to point out).

Instead of issuing a correction, however, Mehlhorn made the exact same incorrect claim four days later on a different website:

We spend a lot more than we used to.

Figure 3. Inflation-Adjusted Cost of a complete K-12 Public Education, and Percent Change in Achievement of 17-Year-Olds, since 1970



America's schools today spend about 2.5 times per pupil what they spent in 1970, notwithstanding a small per-pupil dip since 2008. Ben acknowledges "the fact that K-12 spending has risen in inflation-adjusted dollar value terms over the past 45 years," but then waves that away by saying that "real spending on practically everything has increased in dollar terms since the 1970."

That statement is jarringly untrue.

Over that time period, per-unit prices have plummeted in many areas, including appliances, telecommunications, electronics, computers, televisions, and audio-visual devices. Some sectors have taken advantage: for instance, U.S. military spending has increased only 10 percent since 1970, while dramatically improving its comparative and absolute effectiveness. True, declining costs in some sectors have been offset by price increases in other areas, but this overall mix is called "inflation." By using "inflation-adjusted" dollars, we account for the interplay of cost increases and cost declines. If we ignored inflation, the increase in dollars would be 14 times rather than merely 2.5 times.

Again, and as I pointed out in my first update, I am happy to discuss the importance of real spending increases. But the claim that my true statement was "jarringly untrue" was egregiously wrong and stands uncorrected in multiple places.

3) Comparing research on test score gains and gains in years of completed schooling:

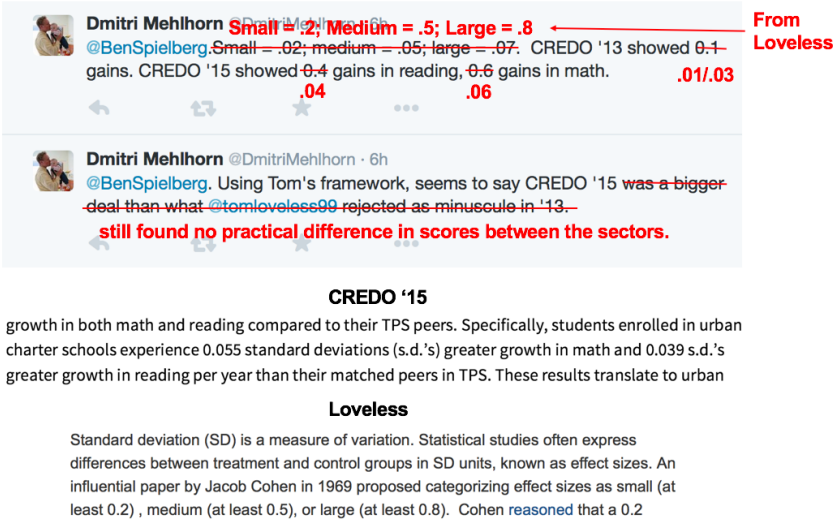
Mehlhorn made the following comparison in his original piece:

> Secondly, what can be done at scale with more money is often much less than what can be done better with existing funds. While any effort at precision will lead to a false sense of certainty, the scale of the difference is clear. For example, the CREDO report showed that urban charter students obtained the equivalent of 40 extra days of math instruction, which would add up to 480 days — or more than 2.6 school years — over the course of 12 years. By comparison, Jackson concludes that a 10 percent increase in funding would result in .44 years of extra schooling for poor children. A zero-cost investment, therefore, would deliver about six times the impact of the $60 billion additional national investment that Jackson's team suggests.

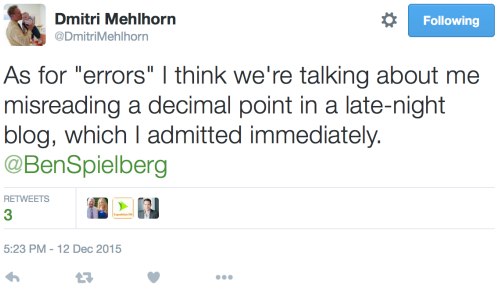
This comparison is totally wrong; I wrote the following in response:

> A new high-quality study by C. Kirabo Jackson, Rucker Johnson, and Claudia Persico comes to the same conclusion. Mehlhorn's article also mentions this study, but misinterprets the results; it mistakenly compares the invalid "months of learning" statistic from the charter school research discussed above (which actually represents data on student test scores) with Jackson et al.'s data on completed years of schooling.
>
> In reality, Jackson et al.'s results are much more striking than most results in education research; the researchers argue in *EducationNext* that, "for low-income children, a 10 percent increase in per-pupil spending each year for all 12 years of public school is associated with roughly 0.5 additional years of completed education, 9.6 percent higher wages, and a 6.1-percentage-point reduction in the annual incidence of adult poverty." While they concede in a follow-up piece that increased school funding won't "eliminate all differences in outcomes by socioeconomic status," they contend "that a 22.7 percent spending increase is large enough to eliminate the average outcome differences between the poor (those with family incomes below twice the poverty line) and the non-poor (those with family incomes above twice the poverty line)."
>
> The researchers' claims here are overstated – they're extrapolations beyond the actual results that, while less misleading than the "months of learning" statistic, are still misguided attempts to help a broader audience understand research findings – but it's important to note that the magnitudes are very large relative to the results in most education studies.

When I posted my piece and explained to Mehlhorn both why the comparison between the two studies is inappropriate and why the months of learning statistic is bogus, Mehlhorn wrote the following tweets, which I annotated and sent back to him along with relevant quotes from the research and argument he referenced:



Based on the fact that the reasoning in his tweets pre-annotation leads to the exact opposite conclusion when the correct statistics are used, I hoped that Mehlhorn would acknowledge that he had misinterpreted the practical significance of the research. Instead, while the numerical mistake is the only one Mehlhorn admits to having made, he insists that it is unimportant and an anomaly:



In both of his follow-up pieces, Mehlhorn just omits the incorrect comparison he originally made (between the CREDO and Jackson et al. results) and the "months of learning" statistic (he also switches his CREDO analysis from aggregate numbers to numbers for student subgroups):

> Ben then takes issue with the growing consensus that charters work, by stating that students in urban charter schools "perform just about as well" as students at district schools. He rests this claim on the fact that, on average, black students in poverty perform eight hundredths of a standard deviations better in math and six-hundredths of a standard deviation better in reading when they are in charter schools, while the numbers for Hispanic students in poverty are, respectively, seven-hundredths and thirty-five hundredths of a standard deviation.
>
> I address the significance of these numbers in a long exchange with another reform skeptic, Mark Weber, on Weber's site. Suffice to say, these average differences matter a lot given that (a) they occur every year, and (b) they are national averages that include jurisdictions that have terrible charter authorizers and terrible charter schools. Additionally, the evidence is very strong that the proliferation of charter schools tends to improve the performance of traditional district schools, perhaps because of healthy competition and the spillover of innovation.
>
> If Ben is sincere in his claim that it matters how money is spent, he should stop nit picking the pro-charter evidence that becomes stronger every few months, and demand that school systems embrace charters before getting new resources.

To be clear, the above paragraphs are an improvement over what Mehlhorn argued in his original piece, though his statement that "the evidence is very strong that the proliferation of charter schools tends to improve the performance of traditional district schools" is not actually supported by any evidence and Mark Weber's response to Mehlhorn, in which he picks Mehlhorn's arguments apart, is definitely worth reading (Mehlhorn's second number for the "Hispanic students in poverty" subgroup is also wrong – it should say "three-and-a-half hundredths"). However, Mehlhorn should have corrected his original piece and explicitly noted in the follow-up pieces how much he got wrong initially.

That's not nit-picking – that's a request for honest, transparent debate.

4) Charter school spending and test scores at charter schools:

After citing the CREDO results above on charter school effectiveness in the aggregate, Mehlhorn wrote the following in his original piece:

> These results, by the way, come
> even though charters spend $1,800 per-pupil less than traditional public schools.
>
> Traditionalists claim that charters succeed by taking the best students or pushing out the worst students. Research since 2009 has empirically rejected these claims. But the most decisive repudiation emerged recently from analysis of the charter sector in New Orleans.
>
> Since the city of New Orleans moved to a charter system, Tulane scholar Doug Harris was able to assess the impact of a system-wide move to charter schools by comparing post-Katrina performance in New Orleans to that of nearby Baton Rouge (which also suffered terrible hurricane damage but did not switch to an all-charter model). Harris wrote of the New Orleans result that "on average student outcomes is quite positive by just about any measure. … We are not aware of any other districts that have made such large improvements in such a short time. The effects are also large compared with other completely different strategies for school improvement, such as class-size reduction and intensive preschool."
>
> The evidence shows that other nations provide high-quality education to their children while spending significantly less money than spent in America – and that charter schools, with little in the way of bureaucracies that typify traditional districts, deliver significantly better results than their counterparts. We even know from the New Orleans experience that charter schools can improve student achievement across an entire system, at more significant levels than expensive interventions such as class size reduction and universal preschool.

I pointed out in my piece that the aggregate spending numbers Mehlhorn cites for the charter sector may be misleading:

> As Baker, Ken Libby, and Kathryn Wiley found in a careful 2012 analysis of charter school and traditional public school spending:
>
> > Comparative spending between the two sectors is mixed, with many high profile charter network schools outspending similar district schools in New York City and Texas, but other charter network schools spending less than similar district schools, particularly in Ohio.
>
> Mehlhorn's counterclaim that charter schools spend significantly less money than traditional public schools likely stems from a 2011 report from the National Center for Education Statistics, but it, like more cursory and flawed studies, may fail to appropriately categorize spending that should be assigned to each type of school. Transportation funding and spending on food services and special education, for example, can be misclassified in such analyses.

I also wrote the following, responding in large part to Mehlhorn's claims about New Orleans:

> There is, of course, important variation in the charter sector; some studies indicate that students in some charter school networks do very well. As Baker, Libby, and Wiley note, however, many of these networks spend substantially more per pupil (sometimes well over 30 percent more) than comparable public schools. Similarly, the test score gains in New Orleans charters that Mehlhorn applauds came with a substantial price tag, a fact that his article conveniently omits. The following excerpt from an interview with researcher Doug Harris is instructive on this point:
>
> > At the beginning New Orleans was spending about $8,000 more per pupil relative to similar districts. In other words, spending didn't quite double, but it came pretty close to doubling in the initial years. And then it converged back to the normal, or close to normal rate. Now they're spending about $1,000 more per pupil than similar districts, whereas before the storm they were spending close to the same as those comparison districts.
>
> Harris doesn't believe the test score gains in New Orleans were entirely a product of increased funding – he finds that explanation unlikely and thinks "every element of the reform package, including the change in spending, probably contributed in some fashion" – but acknowledges that it's possible that increased funding played the primary role. In addition, while Harris thinks there are important lessons to be learned from school reform there, he doubts "you'd see the same effects in other places because the conditions [in New Orleans] were distinctive."

I concluded with my perspective on charter schools and funding:

> Either way, to the extent that best practices in certain successful charter schools drive their results, these practices can likely be replicated in traditional public schools that receive more adequate funding, as research by Roland Fryer suggests. Especially because rapid charter school expansion has often led to harmful side-effects (in New Orleans, the large-scale firing of Black teachers and inattention to community preferences are poignant examples), our efforts are best focused not on promoting charters, but on adequately and equitably funding *all* schools, thus enabling them to implement best practices that may include but are not limited to better teacher training and support, more competitive teacher pay (to facilitate recruitment and retention), reduced class sizes, extended learning time, expanded tutoring availability, and enhanced extracurricular opportunities.

I was hoping Mehlhorn might admit in his follow-ups that he cited New Orleans – a place that saw huge amounts of extra funding – as a core piece of supporting evidence, and that it isn't. I

was also hoping he would admit that his strong suggestion that charters spend less while getting better results doesn't hold up when we look at the charter networks typically held up as models of success. Unfortunately, neither of those things happened. Mehlhorn instead posted a response in both of his follow-ups, shown below, which attempted to change the terms of the debate:

> *The evidence about charters:* The next category of arguments from Ben relates to charter schools. I argue that charter schools deliver better results for urban students in poverty, without spending more money. If true, this establishes yet another independent proof point that school spending is not the primary barrier to educational opportunities for our poorest children. So, is it true?
>
> Well, it seems clear that charters spend less than traditional district schools, or at least do not spend any more (on average). In my first post, I started with an NCES report from 2011 showing that charters spend $1800 less per pupil. Ben cites Baker to rebut that study, but in this case Baker's data manipulations only amount to hand waving that spending in the charter sector varies widely (some charters spend more, some district schools spend less), and that data is hard to obtain (because outside spending such as school fundraisers are not consistently tracked). These narrow points may be true, but it is hard to see how they overcome an aggregate gap of nearly two thousand dollars per pupil.
>
> Ben then notes that some successful charter experiments involve new resources, but this is again not an aggregate number, and Ben ignores the many cases where substantial new resources into districts delivered no results. Ben cites work by Roland Fryer that traditional district schools can replicate some charter practices with more spending, but Fryer's work focuses mostly on practices such as instructional time, data-driven instruction, and cultures of high expectations that have been repeatedly thwarted by unions during negotiations.
>
> In other words, Ben throws a lot at the wall, but nothing sticks to rebut the basic point: Charter don't spend more than traditional districts and their schools.

His original "basic point" seemed to be that charters spend less money than traditional public schools while getting significantly better results. I showed that that point was untrue, that charter schools at which students score well on standardized tests often seem to spend significantly more than other schools, and that public schools that get more money can adopt the useful practices in place at some of the better charters. He responded by making a different point – "charters don't spend more than traditional districts and their schools" – a point which I suspect is true in the aggregate, but which in no way supports his earlier claims about charter schools and school funding.

In other words, this response is extremely misleading; he should have explicitly noted that I rebutted his earlier points (and posted an update to his original piece to note that New Orleans schools received substantial new funding). I would have been happy to discuss how we can make sure schools that receive extra funds use them wisely (despite Mehlhorn's claims to the contrary, unions actually tend to support the best interventions Fryer studied), but that's a different topic entirely.

5) The Bruce Baker/Eric Hanushek debate on school funding research:

This is where Mehlhorn really goes off the rails; his attack on Baker, especially given all of the errors and misleading writing in his own work, is inexcusable. First, he writes the following:

> Twitter, of all things, provided me with news to the contrary in August of 2014. Mark Weber, a sincere reform skeptic and public school teacher in New Jersey, goes by the Twitter handle "Jersey Jazzman," and is a part-time doctoral student at the Rutgers University Graduate School of Education. Weber pointed me to publications written by one of his professors, a fellow named Bruce Baker. Although I had never previously heard of Baker, Weber was not the only person who recommended him. Ben Spielberg, who graduated from my undergraduate alma mater, told me flatly that Baker's "research is legitimate." Spielberg, Weber, and other reform skeptics cited Baker often, and indeed Baker was described by AEI's Rick Hess as the 40th-most cited education scholar in America. Even better, it seemed that Baker was willing to engage folks I knew to be smart and careful, such as Ulrich Boser at the Center for American Progress and Rebecca Sibilia at EdBuild.

Sounds like he didn't know much about Baker before I told him Baker's research was legitimate, right? Wrong. For reference, the Twitter conversation with me that Mehlhorn linked happened on July 31, 2015.

Mehlhorn tells a story in which he is initially wowed by Baker's research. He writes:

> Baker's papers blew me away. They totally reversed the narrative. For instance, Baker pointed me to a 2012 piece he wrote called "Does Money Matter in Education," which concluded that school spending is important and impactful for students.
>
> This conclusions was the opposite of the consensus in academia when I had been a student in the 1990s. How had the prior research been so wrong? What had happened in the previous 15 years? Well, Baker cited Northwestern University's Larry Hedges, who re-reviewed Hanushek's 1986 survey of evidence using "quality control measures" to exclude some studies and change some interpretations. According to Baker, this settled the matter: "by the early 2000s, the cloud of uncertainty conjured by Hanushek in 1986 had largely lifted in the aftermath of the various, more rigorous studies that followed."
>
> I was surprised, but frankly relieved. As I wrote in response to Baker at the time, "Thank heavens. Someone who actually talks evidence. Talk soon." Shortly thereafter, I read another piece from Baker regarding implementation of high-stakes testing, and frankly his analysis was solid. I assumed that this level of analysis was typical of Baker's work, and was further relieved that a high-profile reform skeptic was taking the time to do careful research. As I wrote to him, "Bruce, your facts & analysis R best I've seen on UR side. Wish AFT/NEA pushed you, not smears." I circulated Baker's work to elevate that approach.

The quotes above came on August 9, 2014 and August 11, 2014. But then, according to Mehlhorn, he digs deeper and realizes Baker is not to be trusted:

But wait, something smells fishy

Something was starting to smell fishy

The first clues that something was fishy came as I dove deeper into Baker's body of work. The highly respected Ulrich Boser had written a report on waste and inefficiency in school spending, and Baker had written a rebuttal. Baker's rebuttal was, as I wrote to him, "More strident, less compelling than UR usual." I was being delicate; Baker's rebuttal was full of personal insults and exclamation points. Disappointing for an alleged scholar.

Then, I read a Baker critique of Mathematica policy research regarding the effectiveness of KIPP charter schools. Baker's critique was terrible, a long list of hand-waving attacks that seemed to call into question the very possibility of actual empirical research in education. As I wrote to him, the methodology of his approach seemed like that of climate denialists, whose attacks often are a "kritik" of the very idea of research.

Things got worse still when I started to read Baker's work about teachers' unions, a subject about which I had substantial personal exposure from visiting state legislators in places where unions were active. As I wrote to Baker in response to a blog of his on the subject, his thumb appeared to be on the scale of the internal workings of his models. His methodology on unions was so sloppy it seemed deliberate.

His "less compelling than UR usual" comment comes on August 12, 2014, three days after he apparently first encountered Baker (Mehlhorn seemingly reads very quickly, which may explain his penchant for missing important details). The comment on the Mathematica critique came on August 17, 2014. In between that comment and the comment Mehlhorn made about teachers unions on October 9, 2014, he and I had what I believe to be our first ever interaction – on September 8, 2014. Baker was not mentioned during that conversation, as far as I can tell.

Finally, the Jackson et al. study came out and Mehlhorn went on a fact-finding mission. He had an epiphany:

Yes, I actually believed Bruce Baker was a scholar

In 2015, Rucker Johnson and others published an NBER analysis of the impacts of school spending. The NBER report was broadly sympathetic to Baker's 2012 claims that money can matter, so I read the report with interest.

Wait a minute . . . the 2015 NBER report, entirely focused on the question of "does money matter in education," did not once mention the Bruce Baker publication from 3 years earlier with the title "Does Money Matter in Education?"

That seemed odd. Even more odd, the NBER paper referred to studies from 1995 and 1996 that showed school spending doesn't lead to better results. ***Wait, what?*** Wasn't that the period of time that Baker reviewed, when he wrote that the "cloud of uncertainty" created by Hanushek in 1986 had lifted based on subsequent work? Why didn't Baker mention those 1995 and 1996 studies by other scholars?

With my antennae finally up, I dug into Baker's 2012 claims more fully. As it turns out, Baker omitted so much context from his report that his conclusion borders on outright mendacity. For instance, Baker chooses not to mention that Hanushek wrote several peer-reviewed rebuttals to Hedges' work, including that they engaged in "statistical manipulations … to overturn prevailing conclusions," and that they "misinterpret the implications of their analysis [and,] through a series of analytical choices, systematically bias their results toward the conclusions they are seeking." Baker wrote a conclusion that "uncertainty" created by Hanushek "lifted" after 1986, without even deigning to mention that Hanushek didn't agree? Baker's presentation of this conclusion was so skewed that later scholars on the exact same subject didn't even *mention* Baker's paper?

Judging from a comment on it, the article Mehlhorn mentioned came out on or before July 8, 2015, more than three weeks ahead of the conversation he and I had in which I asserted that Baker's research was legitimate. By his own admission and given his extremely fast reviews of other research, Mehlhorn clearly didn't trust Baker at this point – it is absurd of him to imply, as he does in the piece, that he was an unsuspecting victim of my encouragement to read Baker's work.

What about Mehlhorn's critiques of Baker, which also appeared in his first "rebuttal" to my piece (see below)?

The conclusions of Professor Bruce Baker: Even more than Jackson and his team., Ben relies heavily on articles published by Bruce Baker of Rutgers University's Graduate School of Education. This reliance is common among reform skeptics, as Baker reaches the most anti-reform conclusions to be found within mainstream academia. Particularly cited by Ben is Baker's 2012 editorial published by the Albert Shanker Institute in which he writes that "by the early 2000s, the cloud of uncertainty conjured by Hanushek in 1986 had largely lifted in the aftermath of the various, more rigorous studies that followed." Baker justifies this claim largely by citing Northwestern University's Larry Hedges, who re-reviewed Hanushek's studies "quality control measures." Reading Baker's paper by itself, it is understandable why Ben finds a clear academic consensus that money matters.

The problem is that Baker omits so much that his conclusion borders on outright mendacity. For instance, Baker chooses not to mention that Hanushek wrote several peer-reviewed rebuttals to Hedges' work. One of Hanushek's responses could have been written with Ben in mind: "Hedges, Laine, and Greenwald commit the larger error of asking the wrong question. This problem tends to get lost in their statistical manipulations and their zeal to overturn prevailing conclusions about the effectiveness of pure resource policies in promoting student achievement."

A later paper from Hanushek goes into great detail about how Hedges and company "misinterpret the implications of their analysis [and,] through a series of analytical choices, systematically bias their results toward the conclusions they are seeking." While Hanushek's rebuttal is devastating, the more important point is that Baker simply pretends it does not exist – he paints a story of academic consensus that is entirely false.

In assessing Baker, it is worth noting that serious education researchers tend to not even mention Baker. Jackson and his team, for instance, write an entire paper that "money matters", and don't once mention Baker's 2012 editorial. Rather, they refer to studies from 1995 and 1996 (which Baker ignores) that school spending doesn't lead to better results.

The reason Baker gets so little play in serious education academia is because he writes editorials, not studies. His analyses are designed to achieve his intended results, and he does this by making subjective and one-sided decisions about what to include and what to ignore. [This is a point **Dropout Nation** Editor RiShawn Biddle hit upon four years ago.] This is expected for expert witnesses at trials, but it is disturbing for someone who pretends to be an academic, and is not transparent that he gets paid for reports by parties with a direct financial stake in his outcomes.

This problem was underscored in a 2011 tape-recorded conversation in which Baker said he would play with data, manipulate the questions he asked, and "pull things in and out" of his models "to tell the most compelling story" in exchange for a substantial research grant. This telephone conversation, including Baker's own partially exculpatory comments, appears in full at about the 3-minute mark of this video clip. [Baker offers a rather lengthy explanation and defense of what happened.]

None of this automatically invalidates Baker's conclusions, but most of his research suffers the same kinds of glaring deficiencies I just mentioned regarding his 2012 Shanker Institute paper. Some day, someone may decide to write a point-by-point review of Baker's editorials, but for now the main point is to take his sweeping anti-reform conclusions with a heaping of salt.

Well, for starters, after Mehlhorn wrote that first rebuttal, I noted the following:

> 1) Mehlhorn devotes a lot of space to attacking Bruce Baker for editorializing. Baker certainly does have strong opinions, but I actually think it's nice that he's transparent about his perspective – all researchers have biases, and it's in many ways preferable to know about them upfront. Baker's work is strong and consistent with other recent research. The research Mehlhorn relies on – from Eric Hanushek, a member of the Right-wing Hoover Institution (note that Mehlhorn does not once mention *Hanushek's* affiliation and biases) – is typically much older and a clear outlier (as I explained above).

In his anti-Baker follow-up, however, Mehlhorn still mentioned that Hanushek was at Stanford but once again conveniently forgot to mention his affiliation with the Hoover Institution (since Mehlhorn is, like me, a Stanford alum, he should know that there is a huge difference between "Stanford" and "Hoover"). Mehlhorn also fails to mention that Hanushek's conclusions have been criticized by many other well-respected researchers over the years; here is one example from 2001 and another from 2002 (though as the second link notes, there isn't nearly as much

of a chasm between Hanushek's academic research and other research on the subject as Hanushek's policy advocacy – against increased investments in schools – makes it seem).

Note also that Hanushek accepts large payments to testify against increases in school funding *even when he hasn't analyzed relevant state-level data*.

Mehlhorn must not have read Baker's paper as closely as he said he did, either; his assertion that Baker ignores Hanushek's rebuttals to Hedges (who wrote his paper with Rob Greenwald and Richard Laine when they were all at the University of Chicago) is false. Baker mentions them in a footnote, as shown below:

[16] Greenwald and colleagues explain:

"studies in the universe Hanushek (1989) constructed were assessed for quality. Of the 38 studies, 9 were discarded due to weaknesses identified in the decision rules for inclusion described below. While the

remaining 29 studies were retained, many equations and coefficients failed to satisfy the decision rules we employed. Thus, while more than three quarters of the studies were retained, the number of coefficients from Hanushek's universe was reduced by two thirds." (p. 363)

Greenwald and colleagues further explain that:

"Hanushek's synthesis method, vote counting, consists of categorizing, by significance and direction, the relationships between school resource inputs and student outcomes (including but not limited to achievement). Unfortunately, vote-counting is known to be a rather insensitive procedure for summarizing results. It is now rarely used in areas of empirical research where sophisticated synthesis of research is expected." (p. 362)

Hanushek (1997) provides his rebuttal to some of these arguments, and Hanushek returns to his "uncertainty position:

"The close to 400 studies of student achievement demonstrate that there is not a strong or consistent relationship between student performance and school resources, at least after variations in family inputs are taken into account." (p. 141)

Hanushek, E.A. (1997) Assessing the Effects of School Resources on Student Performance: An update. Educational Evaluation and Policy Analysis 19 (2) 141-164

See also:

Hanushek, Eric A. "Money Might Matter Somewhere: A Response to Hedges, Laine and Greenwald." *Educational Researcher*, May 1994, 23, pp. 5-8.

Here are the other studies Baker cites before stating that "the cloud of uncertainty created by Hanushek in 1986 had largely lifted:"

[18] Wenglinsky, H. (1997) How Money Matters: The effect of school district spending on academic achievement. Sociology of Education 70 (3) 221-237

[19] Taylor. C. (1998) Does Money Matter? An Empirical Study Introducing Resource Costs and Student Needs into Educational Production Function Analysis. In U.S. Department of Education. National Center for Education Statistics. *Developments in School Finance, 1997.*

[20] Baker, B.D. (2001) Can flexible non-linear modeling tell us anything new about educational productivity? *Economics of Education Review* 20 (1) 81-92.

Figlio, D. N. (1999). Functional form and the estimated effects of school resources. *Economics of Education Review*, *18* (2), 242–252.

Dewey, J., Husted, T., Kenny, L. (2000) The ineffectiveness of school inputs: a product of misspecification. Economics of education Review 19 (1) 27-45

Oh, and by the way, those studies from 1995 and 1996 that Baker ostensibly ignored? He didn't. Here's the section, on pages 3 and 4, where he references the 1996 study that Jackson et al. cited:

In short, while family background certainly matters most, schools matter as well. Furthermore, there exist substantive differences in school quality that explain a substantial portion of the variation in student outcomes.

Subsequent studies using alternative data sources to explored the relationship between schooling quality and various outcomes, including the economic rate of return to schooling – e.g., future earnings. For example, David Card and Alan Krueger (1992) studied the relationship between school quality measures, including pupil to teacher ratios and relative teacher pay, on the rate of return to education for men born between 1920 and 1949. Card and Krueger found that men educated in states with higher-quality schools have a higher return to additional years of schooling. Rates of return were also higher for individuals from states with better-educated teachers. [12]

Similarly, Julian Betts (1996) provided an extensive review of the literature that attempts to link measures of schooling quality and adult earnings, including Card and Krueger's study. Betts explains that, while the overall results of such studies were mixed, they were generally positive. More specifically, he pointed to more positive results for studies evaluating the association between district-level spending and earnings, as opposed to those attempting to identify a link between *school*-level resources and earnings, for which results are murkier.[13]

This summary highlights a different aspect of Betts' paper than Jackson et al. chose to, which may be why Mehlhorn missed it, but it is accurate. Back in 1996, Betts found both that "[t]he studies that measure spending by state averages almost always find a positive association between educational expenditures and average earnings" *and* that "when researchers have

attempted to identify the specific components of total educational spending that most influence earnings, most studies found either no link or a positive link that is not robust to changes in specification or subsample." Baker highlighted the former – money matters – while Jackson et al. highlighted the latter – at the time, the literature wasn't clear about *why* money mattered and *how* it could be used productively.

Mehlhorn also asserts that Jackson et al. didn't cite Baker's work in their paper "because he writes editorials, not studies. His analyses are designed to achieve his intended results, and he does this by making subjective and one-sided decisions about what to include and what to ignore."

The more likely explanation seemed to me to be that Baker's review was a review, not original academic research, and that Jackson et al. were citing only research that their findings called into question. I spoke with Jackson and he confirmed this intuition; he respects Baker's work and has no issues with Baker's review (which Baker recently updated).

Finally, Mehlhorn tries to impugn Baker's integrity by citing the work of James O'Keefe, a notoriously dishonest "conservative activist" known for "deliberately misrepresenting" information about the individuals he targets and releasing "selectively edited videotape" (O'Keefe also enjoys breaking the law while pursuing his entrapment schemes). Feel free to read the emails yourself; Baker clearly did not agree to anything unethical.

I believe Mehlhorn should both issue a correction and apologize to Baker for inaccurately maligning his character.