



School of Education, University of Colorado Boulder
Boulder, CO 80309-0249
Telephone: 802-383-0058

NEPC@colorado.edu
<http://nepc.colorado.edu>

RESPONSE OF FRANCESCA LÓPEZ TO REBUTTAL FROM CRPE

NOVEMBER 6, 2014

In September, the National Education Policy Center (NEPC) published a think-tank [review I wrote](#) on a report entitled, “A Meta-Analysis of the Literature on the Effect of Charter Schools on Student Achievement,” authored by Betts and Tang and published by the Center on Reinventing Public Education. My review examined the report, and I took the approach that a person reading the report and the review together would be in a better position to understand strengths and weaknesses than if that person read the report in isolation. While my review includes praise of some elements of the report, there is no question that the review also points out flawed assumptions and other areas of weakness in the analyses and the presentation of those analyses. The authors of the report subsequently wrote a prolonged rebuttal claiming I misrepresent their analysis and essentially reject my criticisms.

The rebuttal takes up 13 pages, which is considerably longer than my review. Yet these pages are largely repetitive and can be addressed relatively briefly. In the absence of sound evidence to counter the issues raised in my review, the rebuttal resorts to lengthy explanations that obscure, misrepresent, or altogether evade my critiques. What seems to most strike readers I’ve spoken with is the rebuttal’s insulting and condescending tone and wording. The next most striking element is the immoderately recurrent use of the term “misleading,” which is somehow repeated no fewer than 50 times in the rebuttal.

Below, I respond to each so-labeled “misleading statement” the report’s authors claim I made in my review—all 26 of them. Overall, my responses make two primary points:

- The report’s authors repeatedly obscure the fact that they exaggerate their findings. In their original report, they present objective evidence of mixed findings but then extrapolate their inferences to support charter schools. Just because the authors are accurate in *some* of their descriptions/statements does not negate the fact that they are misleading in their conclusions.
- The authors seem to contend that they should be above criticism if they can label their approaches as grounded in “gold standards,” “standard practice,” or “fairly standard practice.” When practices are problematic, they should not be upheld simply because someone else is doing it. My task as a reviewer was to help readers understand the strengths and weaknesses of the CRPE report. Part of that task was to attend to salient threats to validity and to caution readers when the authors include statements that outrun their evidence.

One other preliminary point, before turning to specific responses to the rebuttal’s long list. I am alleged by the authors to have insinuated that, because of methodological issues inherent in social science, social scientists should stop research altogether. This is absurd on its face, but I am happy to provide clarification here: *social scientists who ignore details that introduce egregious validity threats (e.g., that generalizing from charter schools that are oversubscribed will introduce bias that favors charter schools) and who make inferences on their analyses that have societal implications, despite their claims of being neutral, should act more responsibly. If unwilling or unable to do so, then it would indeed be beneficial if they stopped producing research.*

What follows is a point-by-point response to the authors’ rebuttal. For each point, I briefly summarize those contentions, but readers are encouraged to read the full 13 pages. The three documents – the original review, the rebuttal, and this response – are available at <http://nepc.colorado.edu/thinktank/review-meta-analysis-effect-charter>. The underlying report is available at <http://www.crpe.org/publications/meta-analysis-literature-effect-charter-schools-student-achievement>.

#1. The authors claim that my statement, “This report attempts to examine whether charter schools have a positive effect on student achievement,” is misleading because: “In statistics we test whether we can maintain the hypothesis of no effect of charter schools. We are equally interested in finding positive or negative results.” It is true that it is the null hypothesis that is tested. It is also true that the report attempts to examine whether charter schools have a positive effect on student achievement. Moreover, it is telling that when the null hypothesis is *not* rejected and no assertion regarding directionality *can* be made, the authors still make statements alluding to directionality (see the next “misleading statement”).

#2. The authors object to my pointing out when they claim positive effects when their own results show those “effects” to not be statistically significant. There is no question that the report includes statements that are written in clear and non-misleading ways. Other statements are more problematic. Just because the authors are accurate in *some* of their descriptions does not negate my assertion that they make “[c]laims of positive effects when they are not statistically significant.” They tested whether a time trend was significant; it was not. They then go on to say it is a positive trend in the original report, and they do it again in their rebuttal: “We estimate a positive trend but it is not statistically significant.” This sentence is misleading. As the authors themselves claim in the first rebuttal above, “In statistics we test whether we can maintain the hypothesis of no effect.” This is called null hypothesis statistical testing (NHST). In NHST, if we reject the null hypothesis, we can say it was positive/negative, higher/lower, etc. If we fail to reject the null hypothesis (what they misleadingly call “maintain”), we *cannot* describe it in the direction that was tested because the test told us there isn’t sufficient support to do that. The authors were unable to reject the null hypothesis, but they call it positive anyway. Including the caveat that it is not significant does not somehow lift them above criticism. Or, to put this in the tone and wording of the authors’ reply, they seem “incapable” of understanding this fundamental flaw in their original report and in their rebuttal. There is extensive literature on NHST. I am astonished they are “seemingly unaware” of it.

#3. My review pointed out that the report shows a “reliance on simple vote-counts from a selected sample of studies,” and the authors rebut this by claiming my statement “insinuates incorrectly that we did not include certain studies arbitrarily.” In fact, my review listed the different methods used in the report, and it does use vote counting in a section, with selected studies. My review doesn’t state or imply that they were arbitrary, but they were indeed selected.

#4. The authors also object to my assertion that the report includes an “unwarranted extrapolation of the available evidence to assert the effectiveness of charter schools.” While my review was clear in stating that the authors were cautious in stating limitations, I also pointed to specific places and evidence showing unwarranted extrapolation. The reply does not rebut the evidence I provided for my assertion of extrapolation.

#5. My report points out that the report “... finds charters are serving students well, particularly in math. This conclusion is overstated; the actual results are not positive in reading and are not significant in high school math; for elementary and middle school math, effect sizes are very small...” The authors contend that their overall presentation of results is not misleading and that I was wrong (in fact, that I “cherry picked” results and “crossed the line between a dispassionate scientific analysis and an impassioned opinion piece”) by pointing out where the authors’ presentation suggested pro-charter results where unwarranted. Once again, just because the authors are accurate in *some* of

their descriptions does not negate my assertion that the authors' conclusions are overstated. I provided examples to support my statement that appear to get lost in the authors' conclusions. They do not rebut my examples, but instead call it "cherry picking." I find it telling that the authors can repeatedly characterize their uneven results as showing that charters "are serving students well" but if I point to problems with that characterization it is somehow I, not them, who have "crossed the line between a dispassionate scientific analysis and an impassioned opinion piece."

#6. I state in my review that the report includes "lottery-based studies, considering them akin to random assignment, but lotteries only exist in charter schools that are much more popular than the comparison public schools from which students are drawn. This limits the study's usefulness in broad comparisons of all charters versus public schools." The rebuttal states, "lottery-based studies are not 'akin' to random assignment. They *are* random assignment studies." The authors are factually wrong. Lottery-based charter assignments are *not* random assignment in the sense of, e.g., random assignment pharmaceutical studies. I detail why this is so in my review, and I would urge the authors to become familiar with the key reason lottery-based charters are *not* random assignment: weights are *allowed*. The authors provided no evidence that the schools in the study did not use weights, thus the distinct possibility exists that various students do not have the same chance of being admitted, and are therefore, not randomly assigned. The authors claim charter schools with lotteries are not more popular than their public school counterparts. Public schools do not turn away students because seats are filled; their assertion that charters do not need to be more popular than their public school counterparts is unsubstantiated. Parents choose a given charter school for a reason – oftentimes because the neighborhood school and other charter school options are less attractive. But beyond that, external validity (generalizing these findings to the broader population of charter schools) requires that over-enrolled charters be representative of charters that aren't over-enrolled. That the authors test for differences does not negate the issues with their erroneous assumptions and flatly incorrect statements about lottery-based studies.

#7. The authors took issue with my critique that their statement, "One conclusion that has come into sharper focus since our prior literature review three years ago is that charter schools in most grade spans are outperforming traditional public schools in boosting math achievement" is an overstatement of their findings. In their rebuttal, they list an increase in the number of significant findings (which is not surprising given the larger sample size), and claim effect sizes were larger without considering confidence intervals around the reported effects. In addition to that, the authors take issue with my critique of their use of the word "positive" in terms of their non-significant trend results, which I have already addressed in #2.

#8. The authors take issue with my finding that their statement, “...we demonstrated that on average charter schools are serving students well, particularly in math” (p. 36) is an overstatement. I explained why this is an overstatement in detail in my review.

#9. The authors argue, “Lopez cites a partial sentence from our conclusion in support of her contention that we overstate the case, and yet it is she who overstates.” The full sentence that I quoted reads, “But there is stronger evidence of outperformance than underperformance, especially in math.” I quoted that full sentence, sans the “[b]ut.” They refer to this as “chopping this sentence in half,” and they attempt to defend this argument by presenting this sentence plus the one preceding it. In either case, they fail to support their contention that they did not overstate their findings. Had the authors just written the preceding sentence (“The overall tenor of our results is that charter schools are in some cases outperforming traditional public schools in terms of students’ reading and math achievement, and in other cases performing similarly or worse”), I would not have raised an issue. To continue with “But there is stronger evidence of outperformance than underperformance, especially in math” is an ideologically grounded overstatement.

#10. The authors claim, “Lopez seriously distorts our work by comparing results from one set of analyses with our conclusions from another section, creating an apples and oranges problem.” The section the authors are alluding to reported results of the meta-analysis. I pointed out examples of their consistent exaggeration. The authors address neither the issue I raise nor the support I offer for my assertion that they overstate findings. Instead, they conclusively claim I am “creating an apples and oranges problem.”

#11. The authors state, “Lopez claims that most of the results are not significant for subgroups.” They claim I neglected to report that a smaller sample contributed to the non-significance, but they missed the point. The fact that there are “far fewer studies by individual race/ethnicity (for the race/ethnicity models virtually none for studies focused on elementary schools alone, middle schools alone, or high schools) or other subgroups” is a *serious* limitation. The authors claim that “This in no way contradicts the findings from the much broader literature that pools all students.” However, the reason ethnicity/race is an important omission is because of the evidence of the segregative effects of charter schools. I was clear in my review in identifying my concern: the authors’ repeated contentions about the supposed effectiveness of charter schools, regardless of the caution they maintained in other sections of their report.

#12. The authors argue, “The claim by Lopez that most of the effects are insignificant in the subgroup analyses is incomplete in a way that misleads. She fails to mention that we conduct several separate analyses in this section, one for race/ethnicity, one for urban school settings, one for special education and one for English Learners.” Once again, the authors miss the point, as I explain in #11. The authors call my numerous examples that

discredit their claims “cherry picking.” The points I raise, however, are made precisely to temper the claims made by the authors. If cherry-picking results in a full basket, perhaps there are too many cherries to be picked.

#13. The authors take issue that I temper their bold claims by stating that the effects they found are “modest.” To support their rebuttal, they explain what an effect of .167 translates to in percentiles, which I argued against in my review in detail. (The authors chose to use the middle school number of .167 over the other effect sizes, ranging from .023 to .10; it was the full range of results that I called “modest.”) Given their reuse of percentiles to make a point, it appears the authors may not have a clear understanding of percentiles: they are *not* interval-level units. An effect of .167 is not large given that it may be negligible when confidence intervals are included. That it translates into a 7 percentile “gain” when percentiles are not interval level units (and confidence bands are not reported) is a continued attempt to mislead by the authors. I detail the issues with the ways the authors present percentiles in my review. (This issue is revisited in #25, below.)

#14. The authors next take issue with the fact I cite different components of their report that were “9 pages apart.” I synthesized the lengthy review (the authors call it “conflating”), and once again, the authors attempt to claim that my point-by-point account of limitations with their report is misleading. Indeed, according to the authors, I am “incapable of understanding” a “distinction” they make. In their original 68-page report, they make many “distinctions.” They appear “incapable of understanding” that the issues I raise concerning “distinctions” is that they were *reoccurring themes* in their report.

#15. The authors next find issue with the following statement: “The authors conclude that ‘charter schools appear to be serving students well, and better in math than in reading’ (p. 47) even though the report finds ‘...that a substantial portion of studies that combine elementary and middle school students do find significantly negative results in both reading and math – 35 percent of reading estimates are significantly negative, and 40 percent of math estimates are significantly negative (p. 47)’.” This is one of the places where I point out that the report overstates conclusions notwithstanding their own clear findings that should give them caution. In their rebuttal, the authors argue that I (in a “badly written paragraph”) “[insinuate] that [they] exaggerate the positive overall math effect while downplaying the percentage of studies that show negative results.” If I understand their argument correctly, they are upset that I connected the two passages with “even though the report finds” instead of their wording: “The caveat here is”. But my point is exactly that the caveat should have reigned in the broader conclusion. They attempt to rebut my claim by elaborating on the sentence, yet they fail to address my critique. The authors’ rebuttal includes, “Wouldn’t one think that if our goal had been to overstate the positive effects of charter schools we would never have chosen to list the result that is the least favorable to charter schools in the text above?” I maintain the

critique from my review: *despite* the evidence that is least favorable to charter schools, the authors claim overall positive effects for charter schools—obscuring the various results they reported. Again, just because they are clear *sometimes* does not mean they do not continuously obscure the very facts they reported.

#16. The authors take issue with the fact that my review included two sentences of commentary on a companion CRPE document that was presented by CRPE as a summary of the Betts & Tang report. As is standard with all NEPC publications, I included an endnote that included the full citation of the summary document, clearly showing an author (“Denice, P.”) other than Betts & Tang. Whether Betts & Tang contributed to, approved, or had nothing to do with the summary document, I did not and do not know.

#17. The next issue the authors have is that I critiqued their presentation and conclusions based on the small body of literature they included in their section entitled, “Outcomes apart from achievement.” The issue I raise with the extrapolation of findings can be found in detail in the review. The sentence from the CRPE report that seems to be the focus here reads as follows, “This literature is obviously very small, but both papers find evidence that charter school attendance is associated with better noncognitive outcomes.” To make such generalizations based on *two* papers (neither of which was apparently peer reviewed) is hardly an examination of the evidence that should be disseminated in a report entitled, “A Meta-Analysis of the Literature on the Effect of Charter Schools on Student Achievement.” The point of the meta-analysis document is to bring together and analyze the research base concerning charter schools. The authors claim that because they are explicit in stating that the body of literature is small, that their claim is *not* an overstatement. As I have mentioned before, just because the authors are clear in their caveats, making assertions about the effects of charter schools with such limited evidence is indeed an overstatement. We are now seeing more and more politicians who offer statements like, “I’m not a scientist and haven’t read the research, but climate change is clearly a hoax.” The caveats do little to transform the ultimate assertion into a responsible statement.

#18. Once again, the authors take issue with my pointing out when they make generalizations on small bodies of work. They state, “Later on, Lopez again takes us to task for our review of the small literature on charter schools and educational attainment.” The sentence from the CRPE report that seems to be the focus here reads as follows, “the general picture that emerges is one suggestive of large positive impacts of charter schools on high school graduation and eventual college enrollment.” The authors argue that their caveat, “It is important to note that this literature is still emerging, and currently covers only a limited number of geographic locations” justifies their use of broad conclusions that favor charter schools. The authors take issue with what they describe as a “broad-brush statement” despite the fact my review points to numerous “broad-brush” assertions made by the authors. Their “broad brush”

assertions have no place *given* the caveats they list. *That* is the issue I raise in the review.

#19. The authors claim I “quarrel” about the inclusion/exclusion of studies as the authors claim, then elaborate on why they excluded KIPP studies but included CREDO studies. I did not “quarrel.” I listed facts from their studies: KIPP was excluded, CREDO was not. This information was, I thought, important to include, but my review did not quarrel with or even critique this decision.

#20. The authors then argue that I “[misunderstand their] analysis of the likely sources of bias in the CREDO studies.” They claim that I incorrectly stated that CREDO studies “introduce biases that favor charter schools.” It was the authors, however, who detail the upward bias (their words, not mine) for propensity score matching due to self-selection on p. 7 of their report. They then state that the approach used in CREDO studies introduces “the same issue as propensity score models: it could be that students who self-select into charter schools are different from students at traditional public schools for unobservable reasons.” That they “believe that the CREDO charter estimates if anything, would be biased toward zero” does not discredit the limitations they raised earlier that they themselves claim are consistent in propensity score studies and CREDO studies. Once again, the consistent theme is that the authors raise limitations and then ignore them.

#21. The authors claim that my review criticizes them for allowing CREDO studies to remain in analyses (I don’t – although I do point out the possible bias issue that the authors themselves raised), and in their rebuttal they claim that although what I stated was true, that “[I] failed to mention” that analyses are re-done. It appears to me that the authors “cherry pick” what they find to be a serious omission. I did not obscure the source of my review, and it was a factual statement regarding the main analysis.

#22. The authors state, “It is simply untrue that we believe that ‘a simple tally of conclusions based on positive and negative results accurately and adequately represents the universe of findings without regard to study size, scope, or significance.’ At this point we have serious concerns about whether Professor Lopez understood the statistical analysis in our report. The main analysis is not a ‘simple tally.’” However, I never claimed the meta-analysis was a simple tally. The authors are referring to the section in the review entitled, “The Report’s Rationale for its Findings and Conclusions.” In the review’s preceding section, I detail the findings of the report’s meta-analysis, making it clear that the meta-analysis itself is much more than a simple tally. Their *rationale*, however, for favoring charter schools points out that some findings favor charter schools and other do not. They then proceed to assert that overall, findings favor charter schools. They do this again in their rebuttal. Accordingly, they engage in what I describe as a “simple tally of conclusions” – a judgment based on something beyond the findings of their meta-analysis. The generous helping of insults found in their misguided

explanation, however, suggests the authors had little else going for them. They also argue that social science considers lottery studies the gold standard. My review explains why the lottery studies are problematic; even if they or others label it a “gold-standard,” I would hope the authors would agree that these studies are far from perfect and are not above criticism.

#23. The authors’ next rebuttal argues, “Lopez frets that charter schools are allowed under federal law to use weighted lotteries to allow minorities a greater chance of attending a charter schools. She fails to document which charter schools actually use weighted lotteries.” To be clear: I did not conduct these studies, use these studies, or defend these studies as the “gold standard.” My role was to help potential readers understand the strengths and limitations of the meta-analysis I reviewed. I would have been extremely remiss if I did not point to the weighted lottery problem as well as other limitations of these lottery studies. Ideally, those limitations would have been highlighted and heeded by the meta-analysis authors, but that was not the case. Given that the authors are responsible for the analyses and inferences drawn, it is the authors who failed to understand the potential bias due to weighted admissions and determine whether this was a fact for the studies they analyzed, as well as inform the inferences they drew.

#24. The authors claim, “On page 6 Lopez argues that it is ‘problematic’ that we would make conclusions when each of the methods used by researchers have potential statistical disadvantages. This is an odd stance to take as it essentially implies that social scientists should stop all research.” They also argue that I am “illogical” when I contend that the problems are particularly acute since “both lottery-based and propensity score matching studies [were] significantly related to the effect size in the meta-analysis for mathematics...interjects systemic bias in the analysis.” I was clear in my review that the authors, for the most part, did make limitations (“potential statistical disadvantages”) clear. That is a strength of their meta-analysis. And, as noted above, my stance was not that social scientists should stop research, but I do encourage more responsible research. Noting a caution sign and then driving full speed ahead is still problematic. Regarding the “illogical” comment, if an analysis carried out finds that the particular kinds of studies are associated with the size of an effect, and both kinds of studies are biased (see #6 and #20 above), how is that “illogical?”

#25. The authors take issue with my critique of their application of percentiles, arguing that explaining effect sizes in percentiles has become “fairly common practice.” Even assuming this is true, every “fairly common practice” is not a good practice. I detail why in my review. The extensive example provided by the authors does not substantiate their poor choice in using percentiles, but is yet another example of how this practice can be – and is – used in order to exaggerate effects that simply are not as large as they claim them to be. The authors also state, “Lopez slips into the habit of labeling the size of estimated charter school effects as ‘small’ and then using [sic] these labels against us.

She is perhaps unaware that the U.S. Department of Education will not allow authors to state that an effect size is large or small.” What the authors appear to be unaware of is that the U.S. Department of Education is not in favor of the “size” labels Cohen (1988) presented in his seminal book because they are not meant for *intervention* studies in education. The fact that charter school studies are not intervention studies notwithstanding, effect sizes described in the context of particular samples is not proscribed. This protest is also odd given the above-quoted statement from the CRPE report: “*the general picture that emerges is one suggestive of **large positive impacts** of charter schools on high school graduation and eventual college enrollment*” (*emphasis added*).

#26. The authors then take issue with my critique of their explanation of effect sizes “across all three years of middle school,” which they present as being additive, in the same increments, across the three years. The authors erroneously assumed I was referring to vertically scaled scores (although, as explained below, they also appear to misunderstand what vertically scaled scores are). They then assert that “...charter school achievement studies do not model gains in achievement, but rather changes in students’ relative standing in the test-score distribution” and that my “statement is completely irrelevant for such studies.” In their rebuttal, the authors attempt to describe vertically scaled scores (incorrectly) and “Z scores.” Regardless of what they think vertically scaled scores are, their rebuttal has absolutely no merit. Whether reported in z units, NCEs, or any other standard score, if there was “expected growth” (i.e., students stayed in the same relative place in the distribution), all things being equal, students would show zero growth from one year to the next because they remained in the same place on the distribution. Vertically scaled scores, contrary to the authors’ representation, are scores from tests that have typically been linked by particular items and/or Item Response Theory so that the tests are representative of each other across grades. Although vertical scales are typically designed to have increasing means across grade levels for ease of interpretability (i.e., parents may not been keen on seeing their child’s score remain the same across time), the comparison of scores across time would remain grounded on where the two scores being compared fall in their relative distribution. Their misunderstanding of vertically scaled scores notwithstanding, the authors’ assumption about effect sizes across time remains flawed. The reason for this is due to changes in variability across grades. I cited one recent example in my review, but the issue of variability across grades stems from seminal scholarly work (Thurstone, 1928).

Thurstone, L. L. (1928). Scale construction with weighted observations. *Journal of Educational Psychology*, 19, 441-453.