**Teach For America's Response to Linda Darling-Hammond et al. Study**
**Letter from Abigail Smith, Vice President of Research and Policy**
**April 18, 2005**

Dear Friends,

You may see coverage of a study about Teach For America that Linda Darling-Hammond, Deborah J. Holtzman, Su Jin Gatlin, and Julian Vasquez Heilig released Friday, April 15, and we want you to be aware of our significant concerns about its methodology and accuracy.

Teach For America is likely the most studied teacher education program in the country, and we welcome research that can advance our work and educational outcomes more broadly. However, we believe it is irresponsible that the Darling-Hammond et al. study was publicly distributed before going through adequate independent review. In this case, there is little doubt that such a review would have uncovered the serious flaws that researchers are telling us undermine the study's conclusions.

The Darling-Hammond et al. study purports to examine the impact of teacher certification on student achievement. It looks at fourth and fifth grade student test scores in Houston from 1997-2002, and concludes that while certified Teach For America corps members (essentially those in their second years) do well against a comparison group of certified teachers, uncertified corps members (essentially first years) do less well than certified teachers. The study compares Teach For America corps members to other teachers across the system rather than with teachers specifically in their schools, and the sample sizes for each analysis are not reported, though in some cases they appear to be as small as two to three Teach For America corps members.

Three years ago, the Carnegie Corporation, Hewlett Foundation, and Smith Richardson Foundation commissioned a study by Mathematica Policy Research to assess Teach For America's impact on student achievement. The nearly $2 million study, which was conducted across six geographic regions, was meant to resolve a long-standing debate between Teach For America and its critics, most visible among them Darling-Hammond herself.

In 2004, Mathematica released its study, which concluded that Teach For America corps members outperform even the veteran and certified teachers in their schools in a statistically significant way. The study also found that Teach For America corps members make 10 percent more progress in a year in math than is typically expected, while slightly exceeding the normal expectation for progress in reading.

Mathematica's study went through an extensive peer review process prior to public release and covered five grade levels in six Teach For America sites including Houston, where Darling-Hammond's latest study was conducted. Most significantly, the study's design called for a randomized trial – considered the gold standard for this type of research. Students in a particular grade were randomly assigned to the classrooms of Teach For America corps members and other teachers in their schools, and the researchers analyzed and compared the gains students made between the beginning and the end of the year. The research community was universally impressed with the sophistication and rigor of the study's design.

In contrast, the Darling-Hammond et al. recently released study reflects only two grade levels in a single Teach For America site, draws conclusions from old data, and appears not even to meet the research standards for its own less rigorous design. Perhaps most concerning, the analysis and conclusions of the Darling-Hammond et al. study were not subjected to rigorous review by other objective researchers or the subjects themselves before being released to the press.

While Teach For America was not given the opportunity to ask questions about the study design, or to view the study before it went to the press, our initial look since the public release has revealed significant flaws in the analysis and methodology. In fact, every researcher with whom we have spoken who has seen this study

has concluded there are problems that could invalidate the conclusions. We summarize several of these problems below.

***There are significant concerns regarding sample sizes, calling into question many of the findings.*** Contrary to standard practice, sample sizes for each analysis are not stated in the paper. This omission makes it difficult to interpret the practical significance of the findings. Furthermore, it raises a question as to whether sample sizes were omitted out of carelessness or out of an effort to draw attention away from embarrassingly small samples. Extrapolating from what information is included regarding samples and from our own internal data, there are numerous cases in which sample sizes seem too small to establish practical significance. Reasonable people would agree, for example, that it would be irresponsible to suggest policy changes based on a sample of two or three individuals – something the Darling-Hammond et al. paper does multiple times. As one critical example, the sample sizes for Teach For America teachers who administered the Spanish language Aprenda math and reading tests are so tiny that it is highly dubious to draw conclusions from these data.

***The conclusions stated in the study are out of proportion with the data presented.*** The study makes assertions as to the lesser effectiveness of Teach For America teachers compared to all other teachers in the district. Yet, once the Aprenda results are omitted, as the tiny sample sizes would demand (see above), a fair reader would conclude from the data that Teach For America teachers are actually at least as effective as the comparison group.

***The study's methodology makes it likely that, in the case of the Teach For America teachers, the effect of teacher experience is misinterpreted as the effect of certification. This misinterpretation undermines the fundamental claim of the study and makes it difficult to interpret many of the study's findings.*** In controlling for teacher experience, the study's statistical model ignores the common understanding that teacher effectiveness does not increase in a linear way with experience – but rather increases steeply in the first couple of years and then begins to level out. At the same time, the analysis does not recognize the fact that in Houston over 95 percent of our first-year corps members are not fully certified, and over 90 percent of our second year teachers are. If both of these factors were taken into account, the analysis – and in turn the conclusions – would look very different. The most elementary error in social science is to confuse correlation with causation, and this study's design has invited that very mistake.

***The study does not make the crucial comparison between the effectiveness of Teach For America teachers and that of other teachers in their schools.*** It is vital in this kind of study to ensure that any differences in impact are not attributable to qualitative differences in the schools that teachers are in. This study did not take the obvious step of comparing teachers within schools. In contrast, the design of the Mathematica study ensured that this "school effect" would not cloud the findings. Given Teach For America's concerted effort to place our corps members in precisely the most challenging and highest-need schools within our placement districts, failure to control for school placement is a serious design flaw – one which effectively punishes any studied group that chooses to focus on low-performing schools.

Academic research holds itself to particular standards for good reason. Studies not conducted or presented in a responsible way can, in the hands of those who do not have the research background to accurately interpret the conclusions, have an adverse impact on policy, which in our case means an adverse impact on outcomes for children and families. The premature public release of the Darling-Hammond et al. study circumvented the very review process that ensures policy-makers will make well-informed decisions. We look forward now to the rigorous analysis of the paper's methodology and conclusions and hope that the policy world will await this process, as well.

Best regards,

Abigail Smith
Vice President, Research and Policy