

Evaluating the Evaluation System: Qualitative Research on a System That Builds Local Program Evaluation Capacity

Kylene Chinsio Shen and Joseph Townsend
Empirical Education Inc.



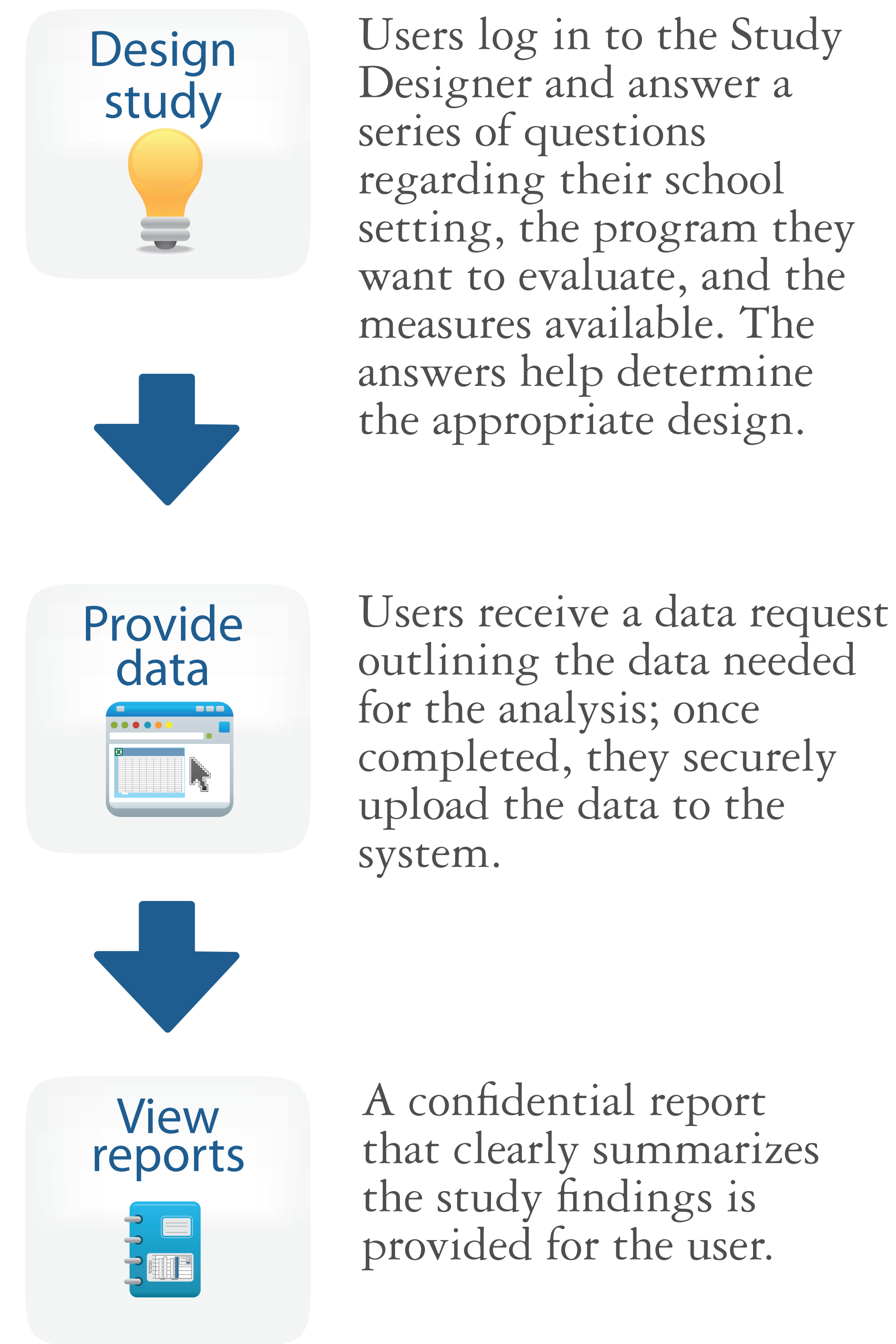
Overview.

Though many districts have data warehouses with capacity for data queries, few have the capacity for conducting the analysis needed to provide evidence of an intervention's effectiveness. In response to this need, we have developed an online program evaluation solution called MeasureResults® that allows district administrators—who may not have a strong working knowledge of statistics or research methodology—to design and implement rigorous local studies and utilize statistical analyses to determine the impact that new programs and interventions have on student achievement.

This idea of bringing statistical analysis tools to school districts reiterates the importance of building capacity for program evaluation at the local level: *by empowering district researchers with tools to conduct local research, districts can expand their capacity for evidence-based decision making and gather evidence easily and efficiently, in a cost-effective way.*

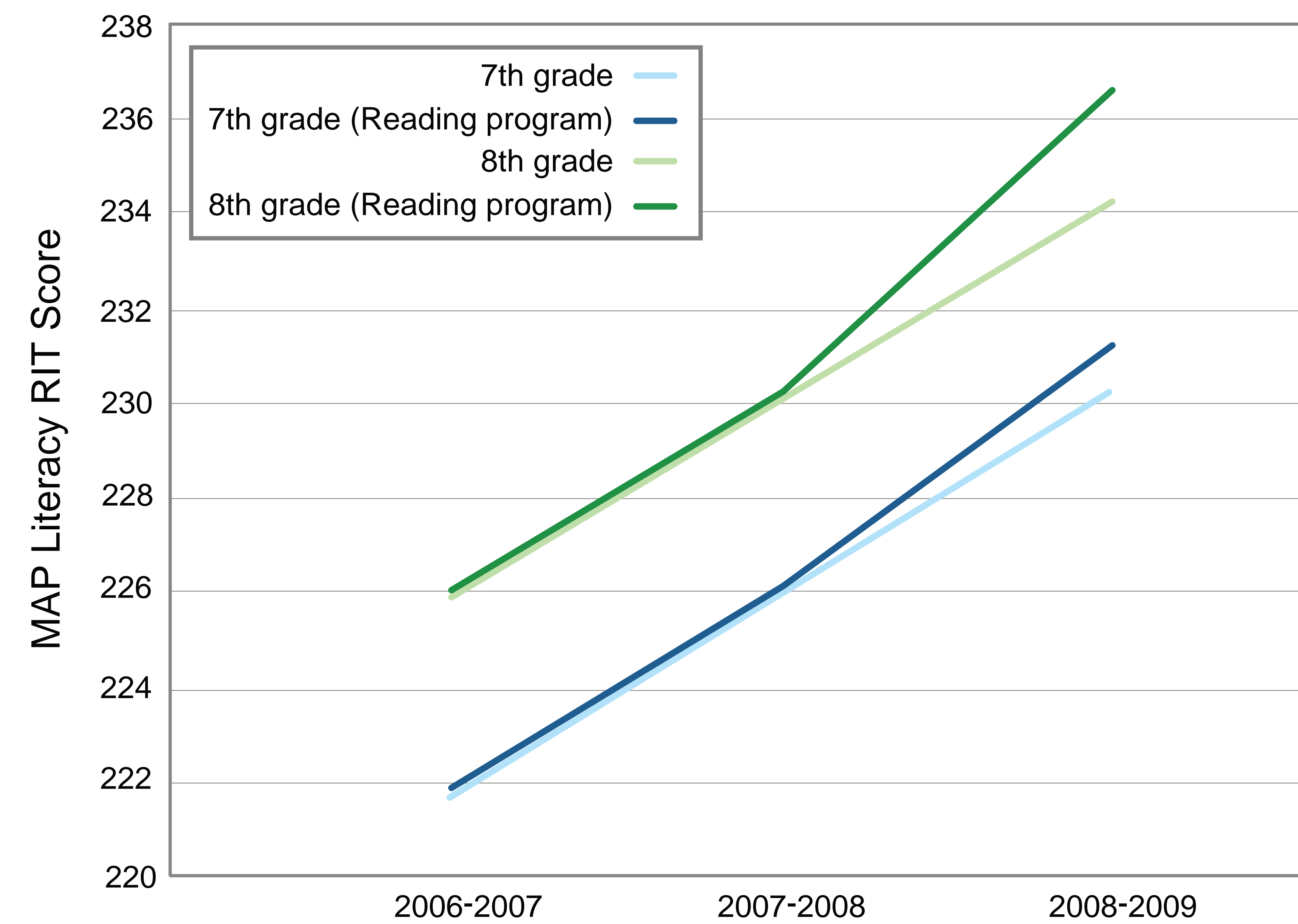
We present the qualitative findings from the first year of formative research on the implementation of MeasureResults, describe the evolution of the tool based on beta testing feedback, and discuss the remaining challenges that the developers are working to address.

MeasureResults® Tool.



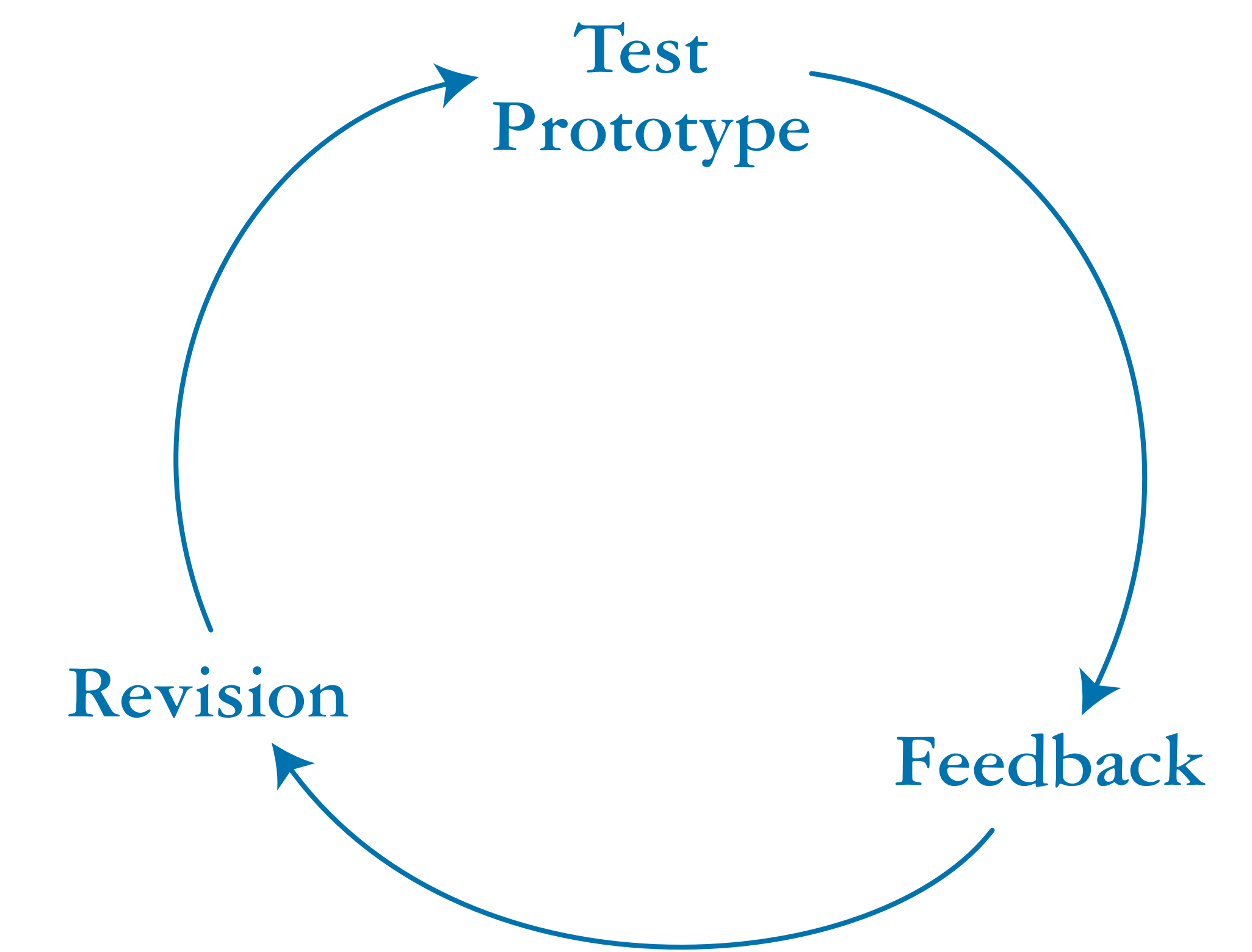
Example.

One district sought to measure the impact of their reading program on middle school achievement, as measured by the NWEA MAP assessment. This was their second time using the system. Feedback from the first study resulted in a more succinct series of design questions, and a clearer graphical representation of the results. The figure below is taken from their report. This report led to a series of new questions, which will be explored using the system in 2011.



Methods.

In our first year of beta implementation, we conducted five district-level program evaluations in three school districts. We followed an iterative process of testing, feedback, and revisions.



All direct feedback from users (collected in the form of surveys, phone interviews, and focus groups) was analyzed for similarities to identify trouble areas or points of confusion within the tool. Lingering questions and concerns about the system are considered areas for future improvement.

Findings: Challenges & Potential Solutions.

Based on user experience and feedback, we identified four main challenges and possible solutions.

Challenge

Identification of treatment and control groups: Varying levels of implementation, collaboration among teaching teams, and “roll out” of new programs to different classrooms at different times

Study Designer dialogue: Assessing the clarity of the questions asked and the completeness of the Study Designer (whether or not users feel that all of their needs and study requirements are communicated)

Data request compliance: System users who design the study may not know what data are available. Also, teacher data, student assessment data, and log usage data from the intervention software system may all be stored in separate databases

Missing and/or “noisy” data: Due to attrition, teacher and student movement within the district, and movement in and out of treatment conditions

Solution

Integrate teacher surveys to determine treatment status: The surveys will take into account the various factors that contribute to intervention usage and will provide a clear treatment designation to be used in the analysis.

Implement automatic summary statements in the Study Designer: Summary statements sum up the main study questions and outcome measures. Develop PD that provide non-researchers with a basic understanding of the different analytic models that may be used, and explain why each piece of information collected is necessary to the process.

Capture data “snapshots” before school systems use the Study Designer: Know what data the user has available upfront to customize the designer accordingly. Build in methods to interface directly with data management systems that are the most widely used among districts.

Develop a sensitivity analysis module: Use imputation and simulation techniques to produce a range of estimates.

Discussion.

Despite the challenges, the three districts found the overall research process to be quite simple from the administrative end. They did have to make adjustments and shifts in staff resources, and the data compilation proved to be more challenging and time-consuming than they had originally expected. However, the staff involved in the research process appreciated the opportunities to participate in a project that they believed added value to their work and contributed to the considerations about student learning. Results of such investigations were then used to inform decisions about instruction, assessment, and program development.

These experiences provide an encouraging outlook for the use and expansion of MeasureResults to a wider audience. We see two main benefits that we expect to be able to replicate in other school systems:

- 1) the ability for districts without a formalized research and evaluation department to continuously conduct rigorous local evaluations, and
- 2) timely use of results from such quick-turnaround studies to inform curricular and program decisions.

Funding for this project is provided by the Institute of Education Sciences, US Department of Education, SBIR contract# ED-08-CO-0046

References:
Lazarev, V. & Smith, R., Jr., (2010). *Quasi-experimental Designs and Models for Studies Using Student-level Data*. Manuscript in preparation.
Nelson, S.R., Leffler, J.C., & Hansen, B.A. (2009). *Toward a Research Agenda for Understanding and Improving the Use of Research Evidence*. Portland, OR: NWREL.