TUSD

1010 E. 10th Street • Tucson, Arizona 85719
OFFICE OF CURRICULUM AND INSTRUCTION (520)225.6285 • fax (520)225.6721

AN EVALUATION OF THE DISCOVER PILOT

August 1, 2015

In Spring 2015, the GATE department initiated a pilot using the 'Discovering Intellectual Strengths and Capabilities while Observing Varied Ethnic Responses' (DISCOVER) assessment developed by Dr. June Maker at the University of Arizona. Created originally in 1987, the Discover assessment is designed to assess students' problem solving abilities as they work through an engaging group of activities and tasks. Utilizing theories such as Gardner's work on Multiple Intelligences, the DISCOVER assessment consists of five problem-sets, or sub-tests that measure spatial artistic, spatial analytic, oral linguistic, written linguistic and mathematical abilities. The ability to look at multiple measures was a key factor in GATE's decision to pilot the DISCOVER assessment with K-1 students. Providing enrichment opportunities for young students can support students' cognitive development as they grow.

The Organization of the Pilot

Selection of Participating classrooms: In order to ensure a diverse population of African American and Latino students, twenty classrooms in 13 Elementary/ K-8 schools were selected from across the District. These classrooms and schools were chosen based on the high number of African American and Latino students (Table 1). Classroom observations were completed between January 26th and March 4th. The following schools participated in the pilot. The numbers in parentheses represent the number of classrooms observed at each site: Bloom (1), Booth- Fickett (2), Erickson (2), Ford (2), Grijalva (1), Holladay (2), Miller (1), Myers-Ganoung (1) ,Naylor (3), Soleng Tom (1), Warren (1), Wheeler (1), Wright (2).

Nine retired elementary teachers –the majority of who were currently substitute teachers for the District - were recruited and trained by a certified DISCOVER trainer over a two week period. Four of the Observers were Hispanic and fluent in Spanish. One Observer served as TEAM Leader, working with the classroom teacher on the transitions between activities in the morning, and guiding the discussions in the afternoon debriefing sessions.

Classroom Observations: All K-1 students present in the selected classes on the scheduled day were assessed with three of the five problem sets of the DISCOVER assessment. Students engaged in activities that were designed to demonstrate **spatial artistic**, **spatial analytic**, and **oral linguistic** abilities. For the Spatial Artistic problem set, students used multiple colored shapes to make pictures, based either on suggested images or in free play. Observers were required to sketch and note any linguistic or general problem-solving strategies. Photographs of the student constructions were taken as well. In Spatial analytic, the students work with Tanagrams to make shapes and complete a puzzle book. This is somewhat similar to the Raven assessment. For Oral linguistic, students were given a set of toys to play with and asked to relate a story (either using the toys or not) to a trained observer. While observers initially scripted the stories, tape recorders were used later in the pilot. This made it easier for the Observers to focus on how a student told their story rather than trying to capture the content. The children were observed while completing these tasks by trained and certified observers who took copious notes, pictures and recordings of student work. An experienced observer watched between four and six students working individually at one time. Observers rotated through the classroom so that each Observer saw about half the students in the class. Completing the three problem sets took up to three hours.

Assessment Findings

Student Characteristics: A total of 443 K-1 students received ratings on these assessments; 123 were kindergarteners and 270 were in first grade. Table 1 provides a breakdown of the students by ethnicity, ELL and IEP status. As the table shows, African

American and Latino students were well-represented at 27% and 41% for Kindergarteners and 21% and 51% for first graders respectively. The percentage of ELL students was also high. Twenty-two percent (22%) of the students were ELL students. ELL students were not only Spanish speaking but there were many refugee students whose first languages were Arabic, Somali and Nepalese. Thirty percent (30%) of the African American students were ELLs and 77 % of the Asian American students. Finally, thirteen percent (13%) of the students was receiving exceptional education services. Thirteen percent (13%) of white students, fifteen percent (15%) of Hispanics, and six percent (6%) of African American students had IEPs. Finally, only eight of the 270 first graders received itinerant GATE services.

Student Ratings: The DISCOVER assessment categorizes students according to demonstrated problem-solving skills. Ratings range from WOW to MAYBE on each problem set based on multiple key factors. These factors look at not only what the students have produced - a drawing, a story, a construction, for example - but the problem-solving strategies students employed to create them. Examples of these strategies include inventing and playing with words, demonstrating concern for proportion, and follow-through. Observers fill out summary forms and checklists for each student and then debrief as a group to determine the criteria for each rating category. The rating criteria can therefore differ across groups. The final ratings for each student in each sub-test were recorded and entered into an Excel spreadsheet.

Table 2 provides a breakdown of the DISCOVER ratings by grade and ethnicity. As the table shows WOW ratings were only given for the Spatial Analytic sub-test. This is not unusual because the assessment relies primarily on quantitative data including time taken, number of puzzles completed, number of puzzle pieces used, and number of clues given. It is therefore easier to assign ratings. A total of thirteen students received a WOW rating with the ethnic breakdown as follows – three White students (4% of all White students), one African American (1% of all African American students), and seven Hispanic students (3% of all Hispanic students). An additional 99 received a DEFINITELY rating, indicating that these students exhibited some of the associated problem-solving

behaviors. This includes twenty White students (23% of all White students), nineteen African American students (18% of all African American students), and 52 Hispanic students (24% of all Hispanic students).

Rating scores were considerably lower for Spatial Artistic and Oral Linguistic. Ten students received a DEFINITELY rating for Spatial Artistic. The ethnic breakdown for Spatial Artistic was one White student (1% of all White students), one African American (1% of all African American students) and five Hispanic students (2% of all Hispanic students). Twenty students received a DEFINITELY rating for Oral Linguistic. This included one White student (1% of all White students), four African American students (4% of all African American students), and thirteen Hispanic students (6% of all Hispanic students). These results indicate that the rating scores did not discriminate with respect to ethnicity.

The Observers attempted to put together criteria for both the Spatial Artistic and Oral Linguistic sub-tests but by the end of the six week pilot period they tended to rely primarily on experience gained from listening to many stories and seeing many constructions. It is extremely difficult therefore to analyze the written documentation to discern how ratings were determined for Spatial Artistic and Oral Linguistic. As a result, no definitive conclusions can be drawn as to whether the rating scores reflect the results of the assessment or an artifact of its implementation (see section below).

Relationship to GATE Testing: One of the primary reasons for piloting the DISCOVER assessment was to assist in identifying students who may not do well on the traditional tests utilized by the GATE department. Only six out of a total of the 270 first graders in the pilot (2% of all first graders) applied for GATE testing in the 2014-2015 school year and these students had all tested the year before. Where GATE test scores were available (fourteen students), there was no relationship between performance on the DISCOVER assessment and the OLSAT/CogAT scores. In fact, the single first grade student who qualified for self-contained placement received no rating higher than a PROBABLY on the three DISCOVER sub-tests. None of the students who met the criteria for GATE services

had a WOW rating. Although the sample is small, this suggests that the DISCOVER assessment is identifying a different type of student than found through the GATE assessments. Although the number of students is negligible (8), there was some relationship with the RAVEN assessment. At least half of the students who performed above the 90th percentile on the RAVEN assessment had at least one DEFINITELY rating. This is not surprising given that the RAVEN is a spatial puzzle assessment.

Identification of Students: The results show that only five percent (5%) of the first grade students participating in the DISCOVER pilot had actually applied for GATE services in the past two years. This indicates that the pilot assessed students who would not normally have applied for GATE testing. Utilizing a selection criteria of a student receiving 2 or more DEFINITELY or higher ratings on the assessment would mean that eight additional students would be identified for GATE itinerant services. This would include seven Hispanic students and one African American student.

Conclusions

Implementation Issues: As with any pilot initiative, particularly with one as complex as the DISCOVER, implementation issues were a factor. Although the two week training was complete, the Observers encountered difficulty with using the various forms/ checklists, and with analyzing the data they collected. There were several revisions over the course of the six weeks and decisions with respect to student ratings changed as the Observers witnessed more students and more classrooms. As indicated, Observers are not considered as proficient without at least 100 hours of observations with the various problem-sets. However, it was clear that by the end of the pilot, the Observers were comfortable and familiar with the activities in the classroom and with the decisions made as a group with respect to student ratings.

Much of the photographic and audio material that was compiled was not available for the final analysis and evaluation. It is therefore difficult to assess the reliability of the student rating scores for Spatial Artistic and Oral Linguistic problem sets. Finally, much of the attention during the Pilot was on addressing implementation issues and as a result there was not enough attention paid to data collection and analysis issues. It is not possible to

draw broad or detailed conclusions based on the summary scores as the criteria was revised throughout the process as the Observers gained experience.

The Strengths of the Pilot: Although the limited quality of the data do not allow for broad conclusions, the results show that rating discriminations could be made across students and that a diverse group of students as defined by ethnicity were rated as WOW or DEFINITELY. In addition, the Pilot reached students who have not been assessed through traditional GATE testing. Only six first grade students of the students tested had applied and tested for GATE in the 2014-2015 school year and all of them had also applied as Kindergarteners.

Specific Recommendations:

1) Invite the eight students who received two or more DEFINITELY or higher ratings for GATE Itinerant services at their school site (subject to space availability). This was a criterion discussed with Helen LePage, the GATE coordinator who oversaw the DISCOVER pilot, prior to her retirement. Based on records none of these students are receiving GATE itinerant services currently, and only one has applied and tested for GATE services in the past year. These students can be followed throughout the course of the year to determine how they differ from students who qualified for GATE itinerant services through the traditional assessments.
2) Consider a further project using the DISCOVER assessment. So much was learned in terms of the mechanics and program implementation that there is little doubt that the focus of another albeit limited project would be spent on better collection and evaluation of outcomes. Although there are budget considerations, many of the expensive items such as training or materials have already been purchased and are available for use.