RFP# DOE-SGA-2012-15
Committee Questions

<u>Offeror Name</u>: Teaching Strategies, LLC
<u>Proposed Assessment Name</u>: *Teaching Strategies GOLD®*

Alignment

Provide evidence of alignment to the current Standards of Learning including a comparative chart of content standards developed by your company to VA content standards for each content area/grade level that your proposed assessment(s) addresses and numbers of items for each standard. If you are planning to develop assessments in response to the contract being awarded, provide the content standards that you will include in your assessment(s) and the number of items you will develop for each standard with a timeline.

The following documents are included with this response as attachments:

- *Alignment of Teaching Strategies GOLD® With Virginia's Foundation Blocks for Early Learning: Comprehensive Standards for Four-Year-Olds*
- *Alignment of Teaching Strategies GOLD® With Virginia Standards of Learning (Kindergarten)*

Student Growth

Provide the rationale for the measure of student growth methodology included in Requirement 3.2 . Also provide the procedures used to validate the measures of growth including statistical processes.

A study was conducted to validate the findings presented in the *Technical Manual for the Teaching Strategies GOLD® Assessment System* (Lambert, Kim, Taylor, & McGee, 2010) as well as to explore the normative data for growth. The research design was cross-sectional, and expected scores were established for children in 3-month age bands. Expected growth for individual children can be inferred from these tables by tracking a child across the age bands as he or she got older. This report supplements the earlier tables by using a longitudinal design and tracking groups of children across a single academic year to observe their growth and development directly.

The growth norms are based on the results of a research study conducted to accomplish five goals:

1. gather evidence that teacher ratings of child developmental status made with the *Teaching Strategies GOLD®* assessment system can be used to track growth across time;
2. on the basis of a nationally representative sample, create norm tables for each scale score that indicate how children of different ages should be expected to score at the beginning of an academic year;

3. create norm tables that illustrate expected growth across an academic year for children of different ages;
4. use the norm sample created for this research study to establish reliability and validity evidence to support the *Teaching Strategies GOLD®* assessment system; and
5. create norm tables that illustrate expected scores at kindergarten entry.

Norm Sample

A norm sample was created from the entire population of children served by teachers who used the *Teaching Strategies GOLD®* assessment system. Children were eligible for inclusion in the sample if their teachers met the following criteria: (a) used *Teaching Strategies GOLD®* for the entire 2011–12 academic year, (b) successfully completed an interrater reliability check conducted by Teaching Strategies, (c) took part in face-to-face *Teaching Strategies GOLD®* training sessions, and (d) successfully completed all online *Teaching Strategies GOLD®* training modules. These criteria helped ensure that the ratings were made by teachers who understood how the tool is intended to be used to assess children's knowledge, skills, and behaviors. With these criteria, 81,375 children were identified as being eligible for inclusion in the norm sample.

From this subpopulation, only children with complete data across three checkpoints were selected for the final norm sample. Many programs record ratings for less than three checkpoints or do not follow a traditional academic-year assessment protocol. Programs typically require teachers to use an October due date for the fall checkpoint, a February due date for the winter checkpoint, and a May due date for the spring checkpoint. However, the particular assessment schedule varies quite substantially among programs. Therefore, a window of acceptable due dates was determined for each assessment checkpoint so that the resulting scores could be interpreted as representing a period of approximately 3 months of growth and development between the checkpoints. Given the admissibility criteria of timely data and complete data, the remaining subpopulation included 69,743 children.

This subpopulation did not represent the exact demographic distribution of children in the United States as measured with respect to ethnicity and race by the 2010 census. Therefore, a random sampling procedure was used to create a norm sample that reflected the most current estimates of the ethnic and racial diversity of the national population of children ages birth to 5 years. Children were included from all regions of the United States, including 34 states, the District of Columbia, and Puerto Rico. The norm sample was 51.3 percent boys and 48.7 percent girls. The primary languages used in the homes of these children were English (77.5 percent), Spanish (17.7 percent), and other languages (4.8 percent). Children with an Individualized Education Program (IEP) or an Individualized Family Service Plan (IFSP) comprised 13.1 percent of the sample. The final norm sample included 54,504 children.

Ratings of Child Growth and Development by 3-Month Age Bands

The norm sample was divided into 3-month bands that were based on the ages of the children at the time of the fall assessment checkpoint. The children were placed into 18 age bands, the youngest band including children aged 3 through 5 months and the oldest band including children aged 54 through 56 months. The *Teaching Strategies GOLD®* assessment system is designed for use with children from birth through kindergarten. However, very few children in the eligible subpopulation were younger than 3 months or older than 56 months at the time of the fall checkpoint. These youngest and oldest children were not included in the final norm sample.

The raw scores are derived by summing across the teacher ratings for all items in each domain of development. Each item was rated by teachers who used a 10-point scale. The ranges of possible raw scores are as follows: Social–Emotional (0–90), Physical (0–50), Language (0–80), Cognitive (0–100), Literacy (0–120), and Mathematics (0–70). The total number of possible raw score points varies by developmental domain because a different number of items is included for each domain.

The scale scores for each domain of development are based on Item Response Theory. They are interval-level scale scores that result from a transformation of the raw scores, are not dependent upon the number of items rated for each domain, and are calibrated to have a mean of 500 and a standard deviation of 100 across the entire age range. The mean of 500 is associated with children at 36 months of age because this is the middle of the intended age-range for the measure. Expected scores are reported for the Social–Emotional, Physical, Language, Cognitive, Literacy, and Mathematics domains of development.

There was an overall trend across the scale scores of approximately 30 scale-score points of growth from the fall to the winter checkpoints for children younger than 30 months at the time of the fall checkpoint. The amount of growth from fall to winter tends to increase to about 35 points for children between 30 and 36 months of age at the time of the fall checkpoint, 40 points for children between 37 and 38 months, and 50 points for those older than 38 months. Additionally, an overall trend across the scale scores of approximately 30 scale-score points of growth between the winter and spring assessment checkpoints for children under 36 months of age at the time of the fall assessment. Children between 36 and 47 months of age at the time of the fall checkpoint tend to make about 35 scale-score points of growth between the winter and spring checkpoints, and children older than 47 months tend to make about 40–45 scale-score points of growth. As expected, the child scores tend to become a little more spread out as the children get older. The standard deviations for any given scale score, at a given checkpoint, tend to be approximately 50 points within the age band for children younger than 30 months at the time of the fall checkpoint and 60 points for those who were 30 months or older.

Reliability and Validity Evidence

A special case of the one-parameter Item Response Theory model, the Rasch partial credit model, was used to examine the measurement properties of the item and scale scores for each of the developmental domains. These analyses were conducted to confirm, on the basis of the current norm sample, the results of the *Technical Manual for the Teaching Strategies GOLD® Assessment System* (Lambert, Kim, Taylor, & McGee, 2010). These analyses were also used to establish that the process used to calibrate the interval-level scale scores for each developmental domain was applicable to the current norm sample. The first step in this process was to examine the model assumption of unidimensionality, i.e., that each scale score is measuring only one underlying construct. This assumption is considered tenable if the model accounts for the majority of the variance in item scores and the first contrast does not account for more than 5 percent of the variance (Bond & Fox, 2007). The results of the principal-components analysis of residuals indicated that the model accounted for over 75 percent of the variance in the item scores for each of the six scale scores (75.7%–84.3%). The first contrast accounted for 5 percent or less of the variance for all scale scores except those in the Physical domain, which was very close to the criterion (6.0 percent). Taken together, these results indicate that the assumption is tenable.

Next, fit statistics were examined to determine whether the item ratings fit the model. Infit (IMSE) and outfit (OMSE) mean-square error statistics of 0.6 to 1.4 are considered acceptable (Bond & Fox, 2007). The criteria were met for every item on the Social–Emotional (IMSE 0.84–1.31, OMSE 0.84–1.30), Physical (IMSE 0.90–1.14, OMSE 0.90–1.13), Language (IMSE 0.83–1.30, OMSE 0.83–1.29), Cognitive (IMSE 0.86–1.15, OMSE 0.86–1.16), and Mathematics (IMSE 0.75–1.14, OMSE 0.75–1.12) scales. The criteria were met for all but one item on the Literacy scale (IMSE 0.76–1.52, OMSE 0.74–1.57). This is strong evidence that the data fit the model across all of the scales.

Item and person reliabilities and Cronbach's alpha reliabilities were examined for each scale score. Values greater than 0.80 are considered acceptable. The results found these coefficients, all of which are at or above 0.90, to be acceptable. Item and person separation indexes were also examined. Values greater than or equal to 3 are considered acceptable, and all of these values met the separation criteria.

Reporting

Provide your best example of a report derived from the assessment's results which illustrates an individual student's growth (not performance).

Teaching Strategies GOLD® Growth Report
Areas of Development: All Areas
Compare to: Color Band Expectations
Checkpoint Periods: Fall 2010/2011, Winter 2010/2011, Spring 2010/2011, Summer 2010/2011, Fall 2011/2012, Winter 2011/2012
Checkpoint Types: Finalized Checkpoint ratings
Age or Class/Grade at Checkpoint Period: Pre-K 4 class/grade (Blue)
Children to Compare: All Children with Ratings in Any Checkpoint Period
Report Level: Organization, Programs, Sites, Classes and Child

Social-Emotional	Color Band		Growth Summary		
	Bottom	Top	Expected Growth	Actual Growth	Type of Growth
Blue Bears (Lead Teacher: Bridgett Scarborough)	589	690	153	160	Expected
Maria Delgado	589	690	166	26	Not Expected
Maria Delgado	589	690	166	170	Expected
David Harvey	589	690	140	141	Expected
David Harvey	589	690	140	-20	Not Expected
Alex Martin	589	690	173	45	Not Expected
Dina Mondsen	589	690	147	153	Expected
Dina Mondsen	589	690	147	13	Not Expected
Matthew Paul	589	690	160	-13	Not Expected
Matthew Paul	589	690	160	178	Expected
Marissa Pierre	589	690	147	49	Not Expected
Marissa Pierre	589	690	147	154	Expected
Courtney Turner	589	690	147	19	Not Expected
Courtney Turner	589	690	147	195	Expected

Technology

For online testing, can portable devices (tablets, iPads®, netbooks) be used with the same fidelity as CPUs/laptops?

Portable devices with an Internet connection and browser can be used to access the *Teaching Strategies GOLD®* online assessment system. The Safari® browser in Apple® portable devices does not support Flash® technology, so a few *Teaching Strategies GOLD®* functions that require Flash® (i.e., uploading files and viewing videos, online professional development modules, and interactive reports) are not available on those devices. However, a separate Apple® application is available that enables file uploading on Apple® devices. With the application, information about student performance can be entered seamlessly into the student's *Teaching Strategies GOLD®* online portfolio.
Because of the many variables involved in testing portable devices (e.g., the ever-changing selection of portable devices on the market, operating systems, Internet connectivity, and cellular data plans) Teaching Strategies has not done exhaustive testing on portable devices.

Can reports be accessed with fidelity from portable devices 24 hours a day, 7 days a week?

Portable devices can be used to access the *Teaching Strategies GOLD®* online assessment system reports 24 hours a day, 7 days a week, 365 days a year. Any upgrades requiring minimal downtime are done late at night and primarily on weekends. The Safari® browser in Apple® portable devices does not support Flash® technology, so viewing the interactive reports, which require Flash® technology, is not possible on those devices. (Interactive reports are the "Widely Held Expectations Report", "Performance and Growth Report", and "Alignment Report" described in Requirement 5.2)

Expand on the technology information provided in Requirement 4.x to include specific requirements about technology infrastructure related to bandwidth, caching capabilities, numbers of concurrent testers, and redundancy of data storage as well as fail-safe protocols during testing windows.

<u>Bandwidth Requirements</u>

A minimum upload and download speed of 300 kbps (kilobits per second) is required for each *Teaching Strategies GOLD®* online user (i.e., each individual administrator or teacher who uses the system). Here are some typical examples:

Internet connection	Maximum number of concurrent users
1.5 Mbps T1	5
6 Mbps DSL	20
10 Mbps Ethernet	34
20 Mbps cable	64
50 Mbps cable	170
100 Mbps fiber	341

RFP# DOE-SGA-2012-15
Committee Questions

Host Site Specifications

Hosting.com hosts *Teaching Strategies GOLD®*. Hosting.com is a privately held Web hosting provider specializing in managed dedicated hosting. As an RSA Managed Security Services Provider (MSSP), Hosting.com provides clients with hosted security solutions based on two-factor authentication. Hosting.com has a 100% uptime guarantee. Teaching Strategies has secured hosting services for all of our online systems (internal and commercial) from Hosting.com since 2001 with virtually no downtime attributed to vendor failure or coding errors.

Network Environment

- Two data centers with collocated servers
- N+1 Redundant UPS Power Supply
- Diesel Generators
- N+1 Redundant HVAC
- Redundant Core Routing
- FE-25 Gas Fire Suppression
- Physically Diverse Fiber Conduits
- Multiple Tier 1 Connectivity Providers
- Central Powerware UPS
- K-Rated transformers
- Redundant routing architecture including hardware failover and Cisco HSRP (Hot Spare Router Protocol)
- Environmentally controlled server room environment HVAC N+1 redundant
- Raised floors in server room
- Optic connections enter the datacenter through physically separate conduits
- Redundant routing architecture, including hardware failover and Cisco HSRP (Hot Spare Router Protocol).
- Redundant uninterrupted power supplies (ups); to handle power outages of at least one hour.
- Backup diesel generator
- System management nightly, including offsite remote two storage arrays third-party backup
- Load-balanced Web server to ensure high availability. If one Web server goes down, the others in the cluster continue serving the site to users.
- Access via the use of redundant T1 circuits
- Stackable switched networks
- Database server(s) (RAID)
- Redundant ISPs
- Optic connections enter the datacenter through physically separate conduits

Data Security

Teaching Strategies GOLD® data is backed up nightly. In the event of a Web server crash, the other servers in the cluster take over. End users will not notice any difference in the functioning of the site. Teaching Strategies and Hosting.com monitor the servers constantly to ensure that they are functioning correctly. If any issues arise during monitoring, we are usually able to resolve the issues seamlessly.

Disaster Recovery

All *Teaching Strategies GOLD®* data is backed up remotely on a nightly basis. In the event of a database corruption or major crash, *Teaching Strategies GOLD®* can be restored to the previous night's backup. In the event of a catastrophic event that destroys the data center where *Teaching Strategies GOLD®* servers are housed, we would be able to rebuild new servers and load them with the secure, remotely backed up data relatively quickly.

During online testing, will remote, "live-time" diagnostic assistance be provided? If so, describe this assistance.

Teaching Strategies GOLD® does not entail student testing, so there is no specific window for online testing. Teachers can access the system at any time. Technical support is available via e-mail and phone, Monday through Thursday, 8:00 a.m. to 7:00 p.m. ET, and Friday, 8:00 a.m. to 6:00 p.m. ET. An online help library includes a *Quick-Start Guide*; help articles; and embedded tips.

What level of local IT support should the division expect in each school/classroom in order to appropriately support successful testing?

Local IT support will be necessary to ensuring that the minimum computer and network system requirements are in place in each school/classroom. Those requirements are specified in the document "GOLD System Requirements," which is included with this response as an attachment.

Availability

For those assessments that are being developed, when will assessments be available for operational use?

The *Teaching Strategies GOLD®* assessment system is fully developed and currently in wide use.