IES Institute of Education Sciences

APPENDIXES
September 2020

Regional Educational Laboratory Southeast
At Florida State University

The Reliability and Consequential Validity of Two Teacher-Administered Student Mathematics Diagnostic Assessments

See https://go.usa.gov/xG4GW for the full report.

Appendix A. Instruments

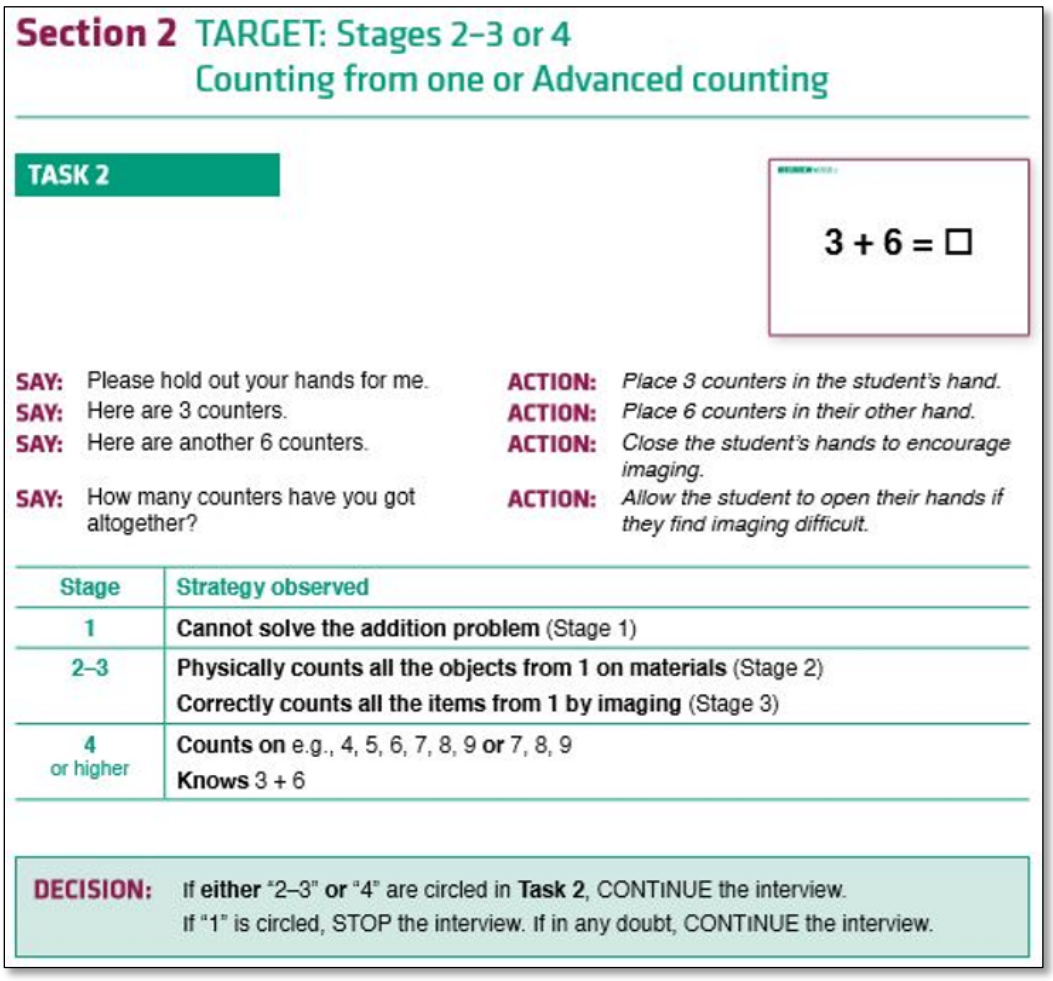
This appendix provides information on the Global Strategy Stage (GloSS) and the Individual Knowledge Assessment of Number (IKAN) assessments, the teacher survey, the focus group protocol, and the fidelity of administration checklists.

Global Strategy Stage assessment

The GloSS assessment provides information on the strategies students use when solving mathematics problems. The GloSS assessment is usually administered before the IKAN, using a one-on-one interview format. The GloSS takes 5–20 minutes to administer. A single form of the assessment was used in this study. The items in the GloSS are not scored as correct or incorrect; rather, students are placed in stages according to the strategies they use to solve problems. Stages range from 0 (one-to-one counting) to 8 (advanced proportional reasoning). The GloSS identifies students' stages in three content area domains: addition and subtraction, multiplication and division, and ratios and proportions.

A sample item from the GloSS is presented in figure A1. For this task the teacher presents the student with the following addition problem (3 + 6 = ___) with a set of counters. Depending on the approach used to solve the problem, the student is placed in one of four Stage Scores (Stage 1 = cannot solve the problem [no strategy]; Stage 2 = physically counts all objects starting with 1; Stage 3 = counts all objects from imaging; Stage 4 = counts on [for example, 4, 5, 6, 7, 8, 9 or 7, 8, 9] or knows 3 + 6 = 9).

Figure A1. Sample item from the Global Strategy Stage administration guide



Section 2 TARGET: Stages 2–3 or 4
Counting from one or Advanced counting

TASK 2

$3 + 6 = \square$

SAY:	Please hold out your hands for me.	**ACTION:**	*Place 3 counters in the student's hand.*
SAY:	Here are 3 counters.	**ACTION:**	*Place 6 counters in their other hand.*
SAY:	Here are another 6 counters.	**ACTION:**	*Close the student's hands to encourage imaging.*
SAY:	How many counters have you got altogether?	**ACTION:**	*Allow the student to open their hands if they find imaging difficult.*

Stage	Strategy observed
1	**Cannot solve the addition problem** (Stage 1)
2–3	**Physically counts all the objects from 1 on materials** (Stage 2) **Correctly counts all the items from 1 by imaging** (Stage 3)
4 or higher	**Counts on** e.g., 4, 5, 6, 7, 8, 9 **or** 7, 8, 9 **Knows** 3 + 6

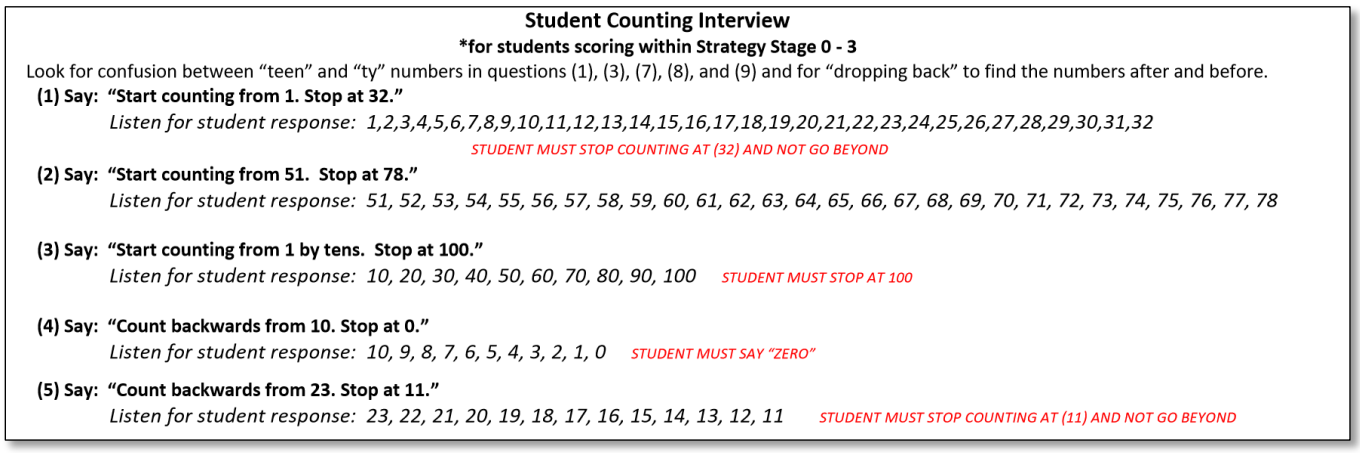
DECISION: If **either** "2–3" **or** "4" are circled in **Task 2**, CONTINUE the interview.
If "1" is circled, STOP the interview. If in any doubt, CONTINUE the interview.

Source: New Zealand Ministry of Education, 2012.

Individual Knowledge Assessment of Number assessment

The IKAN provides teachers with information on students' Number Knowledge Stages across five increasingly abstract domains of arithmetic, including simple whole-number sequencing, multidigit operations, and operations with decimals and fractions. The IKAN assessment is available in two formats, the Counting Interview and the Written Assessment. Students whose GloSS Stage Scores is 0–3 receive the Counting Interview. Sample items from the Counting Interview are in figure A2. In addition, students are shown a series of 12 number cards and are asked, "What is this number?", "What number comes after?", and "What number comes before?" The IKAN Counting Interview takes 5–15 minutes to administer. A single form of the assessment was used in this study.

Figure A2. Sample item from the Counting Interview of the Individual Knowledge Assessment of Number



Student Counting Interview

***for students scoring within Strategy Stage 0 - 3**

Look for confusion between "teen" and "ty" numbers in questions (1), (3), (7), (8), and (9) and for "dropping back" to find the numbers after and before.

(1) Say: "Start counting from 1. Stop at 32."

Listen for student response: 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32

STUDENT MUST STOP COUNTING AT (32) AND NOT GO BEYOND

(2) Say: "Start counting from 51. Stop at 78."

Listen for student response: 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78

(3) Say: "Start counting from 1 by tens. Stop at 100."

Listen for student response: 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 *STUDENT MUST STOP AT 100*

(4) Say: "Count backwards from 10. Stop at 0."

Listen for student response: 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0 *STUDENT MUST SAY "ZERO"*

(5) Say: "Count backwards from 23. Stop at 11."

Listen for student response: 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11 *STUDENT MUST STOP COUNTING AT (11) AND NOT GO BEYOND*

Source: New Zealand Ministry of Education, 2011.

The second format is the IKAN Written Assessment, which is available in both a New Zealand format and an American English format. Because Georgia uses the New Zealand format that format was used for this study. A single form of the assessment was used in this study. Students whose GloSS Stage Score is 4–8 receive the Written Assessment. Each item on this six-minute test is timed. The test includes 40 items, divided across five Stage Scores (4–8). Students are required to respond to each item within the allocated time at a rapid pace. Sample items and answers for the Written Assessment (Form 1) are in table A1. Responses are scored as correct or incorrect. Students receive a Knowledge Stage Score (4, 5, 6, 7, or 8) based on the last stage for which they answered all items correctly. Students unable to answer all of the items correctly for a Knowledge Stage Score of 4 are considered to have a Stage Score of 0–3. Four parallel forms of the IKAN Written Assessment are available; teachers in this study all used Form 1.

Table A1. Sample items and answers for the Written Assessment (Form 1) of the Individual Knowledge Assessment of Number

Item	Part 1 (Stage 4)	Part 2 (Stage 5)	Part 3 (Stage 6)	Part 4 (Stage 7)	Part 5 (Stage 8)
1	What number is one more than 49?	What number is one more 599?	What number is one more 439,999?	Which decimal is the biggest, 0.639, 0.9, 0.84?	Which fraction is the biggest, $^{3}/_{4}$, $^{73}/_{100}$, $^{7}/_{10}$
	50	**600**	**440,000**	**0.9**	**$^{3}/_{4}$**
2	What number is one less than 30?	What number is one less than 1000?	What number is one less than 801,000?	Which decimal is the smallest: 2.4, 2.71, 2.084?	Which is the smallest? $^{2}/_{3}$, 0.6, or 70%
	29	**999**	**800,999**	**2.084**	**0.6**
3	Write the fraction for one half.	Write the fraction for five quarters.	Write these fractions in order of size, smallest to biggest: $^{1}/_{5}$, $^{1}/_{7}$, $^{1}/_{6}$	Which number is the same as $^{3}/_{5}$? $^{5}/_{3}$, $^{12}/_{20}$, $1\ ^{2}/_{3}$, $^{4}/_{6}$	How many hundredths are in all of 6.073?
	$^{1}/_{2}$	**$^{5}/_{4}$**	**$^{1}/_{7}$, $^{1}/_{6}$, $^{1}/_{5}$**	**$^{12}/_{20}$**	**607 or 607.3**
4	Write the fraction for one fifth.	Write these fractions in order of size, smallest to biggest: $^{3}/_{4}$, $^{1}/_{4}$, $^{2}/_{4}$	Write 4 and $^{1}/_{5}$ as a fraction.	Which fraction is the smallest: $^{3}/_{8}$, $^{4}/_{10}$, $^{1}/_{3}$?	What number is halfway between 4.8 and 4.7?
	$^{1}/_{5}$	**$^{1}/_{4}$, $^{2}/_{4}$, $^{3}/_{4}$**	**$^{21}/_{5}$**	**$^{1}/_{3}$**	**4.75**
5	How many tens are in 80?	How many tens are in all of the number 832?	How many hundreds are in all of this number, 53,605?	Round the following decimal to the nearest tenth: 6.49	What is the simplest fraction for 80%?
	8	**83 or 83.2**	**536 or 536.05**	**6.5**	**$^{4}/_{5}$**
6	What is the number for nine groups of ten?	What is the number for 49 groups of ten?	How many tenths are in all of this number, 5.8?	How many thousands are in all of 6,457 894?	What is 1.3 written as a percentage?
	90	**490**	**58**	**6,457 or 6457.894**	**130%**
7	7 + 7 =?	7 + 9 =?	15 – 8 = ?	63 ÷ 9 = ?	What is the least common multiple of 6 and 9?
	14	**16**	**7**	**7**	**18**
8	Half of 18 is ?	5 x 7 = ?	6 x 7 = ?	What number divided by 7 gives 6?	What is the highest common factor of 36 and 48?
	9	**35**	**42**	**42**	**12**

Source: New Zealand Ministry of Education, 2011.

Previous research on the reliability and validity of the Global Strategy Stage and Individual Knowledge Assessment of Number assessments

The study team searched approximately 120 publications for reliability and validity information relating to the GloSS and IKAN assessments and identified three studies as relevant. The only psychometric study of the GloSS that was found used a Rasch model to ensure that parallel forms of this individually administered moderate-inference assessment were roughly equivalent in total test difficulty (Neill et al., 2011). The study did not calculate any type of reliability. When asked whether reliability data were available, Andrew Tagg (NZ Maths, New Zealand Ministry of Education) indicated that only minimal reliability data were available (personal communication, September 24, 2017).

The psychometric study also found that the GloSS appeared to use the same construct as the Progressive Achievement Test of Mathematics, a widely administered standardized test used in New Zealand (Neill et al., 2011). This finding provided some evidence of construct validity and reasonable concurrent validity for the GloSS in relationship to the standardized mathematics achievement test. Thomas and Tagg (2006) traced patterns of

achievement in a longitudinal study conducted in New Zealand, demonstrating that the GloSS is sensitive to student growth in mathematics as a result of instruction, another key aspect of construct validity.

A study by Thomas et al. (2006) addressed criterion validity, contrasting 156 teacher judgments of students' mathematics understanding with the scores determined by researchers administering the GloSS. Although the average agreement between teachers and researchers was 81 percent, only one-third of the participating teachers used the GloSS; the rest used other assessments of numeracy such as the Progressive Achievement Test of Mathematics. No separate analyses were conducted for cases where both the teacher and the researcher assessed the same student at the same time. Thus, this study was deemed to be of limited relevance for establishing criterion validity or interassessor reliability.

Teacher survey

The teacher survey was adapted from the material developed by Gerald Tindal, Daniel Anderson, and colleagues at the University of Oregon (Oregon Department of Education, 2017), based in part on Roach et al. (2007). During the development phase of the current study, an 18-item version of the survey using Google Forms was piloted with a small sample of four teachers to ensure that the items were easily comprehensible. Three additional Likert scale questions were added (two of the added items were about the assessment training), as well as one final open-ended question. The final survey, administered in April 2019, included 22 items: 21 Likert scale items and 1 general open-ended comment question. Items were posed separately for the GloSS and IKAN (table A2).

Table A2. Items from the teacher survey

Based on your experiences using the IKAN and GloSS, together, indicate your level of agreement.
1. Data from the IKAN and GloSS are more useful than screening data from assessments used in our school (e.g., STAR, Easy CBM, AIMSweb) for guiding decisions about which students require intervention.
2. KAN and GloSS data help me address struggling students' instructional needs in my day-to-day teaching.
3. Data from the IKAN are more useful than the data from the GloSS.
4. I prefer using mathematics assessment tools **other** than the IKAN and GloSS for guiding instructional decisions.
5. The time that it takes to individually administer the IKAN and GloSS to each student is worth it because I am able to learn more about how each student reasons and thinks about mathematics.
Based on your experiences using the GloSS, rate the following items in terms of their usefulness. I find the data on strategy use from GloSS useful…
6. …for identifying skills and concepts in which students are weak.
7. …for placement within the RtI or multi-tiered system of support (MTSS).
8. …for modifying instruction for my mathematics class.
9. Overall, I find the data on strategy use from GloSS useful.
Based on your experiences using the GloSS, indicate your level of agreement.
10. Administering the GloSS is a good use of my time.
11. Data on strategy use from the GloSS are useful.
12. I have enough time to administer the GloSS to all the students in my class.
Based on your experiences using the IKAN, rate the following items in terms of their usefulness. I find the data on strategy use from the IKAN useful…
13. ...for identifying skills and concepts in which students are weak.
14. …for placement within the RtI or multi-tiered system of support (MTSS).
15. …for modifying instruction for my mathematics class.
16. Overall, I find the data from the IKAN useful.
Based on your experiences using the IKAN, indicate your level of agreement.
17. Administering the IKAN is a good use of my time.
18. Data on knowledge level from the IKAN are useful.
19. I have enough time to administer the IKAN to all the students in my class.

Based on your attendance at the IKAN and GloSS training/professional development in February, indicate your level of agreement.

20. The training/professional development I received adequately prepared me to properly administer the IKAN and GloSS to my students.
21. The training/professional development I received on the IKAN and GloSS was adequate in helping me identify developmental progressions of students' number knowledge and their ability to solve problems involving operations (e.g., addition, subtraction, etc.).

GloSS is Global Strategy Stage. IKAN is Individual Knowledge Assessment of Number. STAR, Easy CBM, and AIMSweb are assessments used to identify which students are struggling with mathematics. RtI is response to intervention.
Source: Adapted by the authors from material developed by Gerald Tindal, Daniel Anderson, and colleagues at the University of Oregon (Oregon Department of Education, 2017), based in part on Roach et al. (2007).

Focus group protocol

The study team developed a protocol to promote discussions during the focus groups (box A1). Overall, teachers reported in the surveys that the assessments were useful/somewhat useful. The protocol reflects that perspective. The team also developed a focus group script to ensure consistency across groups (box A2).

Box A1. Focus group protocol

Directions. You all have indicated that you find the IKAN and/or GloSS assessments to be useful/somewhat useful. We would like to better understand what features you like best, and why you find them useful.

1. Let's talk about how you used the IKAN and GloSS in your class. Can you provide specific examples for how you used the data from the assessments in your classroom?
   Potential probes:
   a. Has anyone used it in another way?
   b. What types of useful information did the IKAN and GloSS produce to inform changes to your mathematics instruction?
   c. Can anyone give us an example of how the information on *student Strategy Stages* from the GloSS impacted your teaching?
2. All of you indicated that you find the assessments useful/somewhat useful. Could you describe the most helpful aspect of the IKAN and GloSS? Feel free to provide examples...
   Potential probes:
   a. Do you find both the IKAN and GloSS equally useful?
3. All of you indicated that you find the assessments useful/somewhat useful. But there are always things that can be improved. If you could change something to improve the assessments, what would it be?

Source: Authors' creation.

Box A2. Focus group script

1. Welcome the participants.
2. Have participants introduce themselves and have an ice-breaker activity.
3. Describe the purpose:
   a. We have been asked by the Georgia Department of Education to conduct these focus groups to learn more about how the IKAN and GloSS instruments are being used in the state in select districts and to what extent teachers find them useful. They are interested in knowing more about these instruments so they can make more informed decisions about their use within their multi-tiered system of support in mathematics.
   b. We need your input and want you to share your honest and open thoughts with us.
4. Ground rules:
   a. We want you to do the talking. We'd like everyone to participate so I may call on you if I haven't heard from you in a while.
   b. There are no right or wrong answers—every person's experiences and opinions are important. Speak up whether you agree or disagree. We want to hear a wide range of opinions.

c. What is said in this room stays here. We want folks to feel comfortable sharing when sensitive issues come up.
d. We will be tape recording the group to capture everything you have to say. We don't identify anyone by name in our report. You will remain anonymous.

5. Obtain permission to record.

Source: Authors' creation.

Fidelity of administration checklists

The fidelity of administration addresses whether the assessments were implemented as intended, which would strengthen the interpretation of the reliability findings. The study team developed and used fidelity checklists to examine administration of the GloSS, the IKAN Counting Interview, and the IKAN Written Assessment. Each procedure on the checklist was marked as observed, not observed, or not applicable. The checklists included items such as "The examiner uses the decision rules correctly to decide whether to continue with the next tasks in the Addition and Subtraction section."; "The examiner always asks the student to explain his or her strategies."; "The examiner uses the number cards as prescribed in the examiner's manual."; and "The examiner correctly identifies the Stage Score for overall Number Knowledge Stage Score."

References

Neill, A., Lawes, E., Robertson, S., & Darr, C. (2011). *GloSS psychometric study: Report to the Ministry of Education*. New Zealand Ministry of Education.

New Zealand Ministry of Education. (2011). *Individual Knowledge Assessment for Numeracy (IKAN).*

New Zealand Ministry of Education. (2012). *Global Strategy Stage (GloSS).*

Oregon Department of Education (2017). *2016–17 Technical report: Oregon's alternate assessment system.* http://www.oregon.gov/ode/educator-resources/assessment/AltAssessment/Documents/techreport.pdf.

Roach, A. T., Elliott, S. N., & Berndt, S. (2007). Teacher perceptions and the consequential validity of an alternate assessment for students with significant cognitive disabilities. *Journal of Disability Policy Studies, 18*(3), 168–175. https://doi.org/10.1177/10442073070180030501.

Thomas, G., & Tagg, A. (2006). Numeracy development project longitudinal study: Patterns of achievement. In New Zealand Ministry of Education (Ed.), *Findings from the New Zealand Numeracy Development Projects 2005* (pp. 22–33). Learning Media Limited. https://nzmaths.co.nz/sites/default/files/Numeracy/References/Comp05/NumeracyCompendium05.pdf.

Thomas, G., Tagg, A., & Ward, J. (2006). Numeracy assessment: How reliable are teachers' judgments? In New Zealand Ministry of Education (Ed.), *Findings from the New Zealand Numeracy Development Projects 2005* (pp. 91–102). Learning Media Limited. https://nzmaths.co.nz/sites/default/files/Numeracy/References/Comp05/NumeracyCompendium05.pdf.

Appendix B. Methods

This appendix includes details on the sample, data collection, and analyses.

Sample

A combined total of 30 grade 1 and grade 3 teachers participated in the study, 20 from Jefferson County, 6 from Fannin County, and 4 from Walker County, Georgia (table B1). Of the 30 teachers, 6 were on special assignment as mathematics instructional coaches (4 from Jefferson County and 2 from Fannin County). Most teachers were women and had more than 10 years of teaching experience.

Table B1. Demographic characteristics of participating teachers, 2019

Teacher characteristic	Number of teachers (n = 30)
District	
Fannin County	6
Jefferson County	20
Walker County	4
Female	29
Race/ethnicity	
Black	8
White	22
Position in the district	
Elementary teacher	24
Instructional coach	6
Education level[a]	
Bachelor's	9
Master's	13
Post-master's	7
Years of teaching experience	
None	2
1–10	9
11–20	12
21 or more	7
Years of teaching special education	
None	28
1–3	2
Years of teaching in current school	
1	6
2–5	7
6–20	12
21 or more	5

a. Does not sum to 30 because one teacher did not report education level.
Source: Authors' analysis of primary data collected for the study in 2019.

With parent or guardian consent, 32 grade 1 students and 28 grade 3 students participated in the study. The sample was evenly distributed between girls and boys, and most students were either Black or White. The demographic characteristics of the 60 participating students are presented in table B2.

Table B2. Characteristics of participating students, 2019

Characteristic	Percent of student sample
Grade	
1	53
3	47
Gender	
Female	50
Male	50
Race/ethnicity[a]	
Black	48
Hispanic	10
White	43
Declined to respond	7

a. Does not sum to 100 because five students selected multiple categories.
Source: Authors' analysis of primary data collected for the study in 2019.

There was variability across the three districts represented in the sample in race/ethnicity, eligibility for the national school lunch program, and students' mathematics proficiency (based on proficiency level on the Georgia Milestones Assessment in mathematics for the 2017/18 school year). Characteristics of the three participating school districts and Georgia statewide are described in table B3.

Table B3. Characteristics of participating districts, 2017/18

		District		
Characteristic	Georgia	Fannin County	Jefferson County	Walker County
Elementary school locale[a]				
Suburban	na	0	0	4
Town	na	0	1	0
Rural	na	3	2	6
Total teachers				
Grade 1	na	11	9	38
Grade 3	na	11	11	38
Percent of students				
Eligible for the national school lunch program	61	60	100	74
Asian	4	1	1	0
Black	37	0	68	6
Hispanic	16	5	6	3
White	40	92	23	85
Other	3	2	2	6
Percent of students proficient in grade 3 mathematics[b]	34	48	28	31

na is not applicable.
a. Suburban is locale outside a principal city and inside an urbanized area. Town is locale inside an urban cluster that is 10–35 miles from an urbanized area. Rural is locale that is 5–25 miles from an urbanized area.
b. Based on the percentage of grade 3 students who were assessed as Proficient Learners on the Georgia Milestones Assessment in mathematics for the 2017/18 school year.
Source: Data on mathematics proficiency, eligibility for the national school lunch program, and race/ethnicity data are from The Governor's Office of Student Achievement (2020). Data on the number of elementary schools and teachers are from Fannin County School District (2020); Jefferson County School Board of Education (2020); Walker County School Systems (2020). Data on locales are from U.S. Department of Education (2020).

Data collection

GloSS and IKAN assessments. Two different teachers tested 60 students on two occasions within a one-week period using the GloSS and IKAN assessments, beginning with the GloSS, as the developers intended. Teachers were randomly assigned to both time 1 and time 2, with an even distribution to the extent possible so that the student's classroom teacher was not always the first teacher. Teachers were paired with 1.4 other teachers on average (ranging from 1 to 2 other teachers) in the same school. Assessments were administered in March 2019. The study team determined that 60 students were enough to calculate reliable interassessor reliability percentages using either exact agreement (IKAN) or plus-or-minus-one agreement (GloSS).[1]

Fidelity of administration. Two members of the study team conducted fidelity checks of 30 of the 120 testing sessions (25 percent) across Fannin County and Jefferson County School Districts. The teachers were not told which testing sessions would be checked for fidelity. Walker County School District declined to allow the observations needed to determine fidelity of administration.

Teacher survey. Data from the teacher survey were collected in April 2019 from the 30 teachers using Google Forms.

Focus groups. A subsample of 24 teachers participated in focus groups in May 2019. Teachers volunteered for the focus groups depending on their availability and interest in sharing more about their experience with the assessments. Three focus groups were conducted: two in Jefferson County and one in Fannin County. Each focus group included six to nine teachers and lasted about 1.5 hours. Teachers received the protocol before the focus group to prepare for questions that requested examples. Two members of the study team facilitated each focus group. All focus groups were audio-recorded and transcribed.

Analysis

Quantitative analysis. Interassessor reliability for the GloSS and IKAN was calculated using the percentage agreement formula $\left(\frac{Total\ number\ of\ agreements}{Total\ number\ of\ agreements\ and\ disagreements} \times 100\right)$ for exact agreement and plus-or-minus-one agreement. Because the GloSS scoring system requires many sophisticated judgments, using agreement within one Stage Score (plus or minus one) to calculate interassessor reliability was appropriate. Exact agreement was used to calculate interassessor reliability for the IKAN, which is strictly procedural in nature and therefore has a less subjective scoring process than the GloSS.

Fidelity of administration was calculated as the percentage of total procedures implemented, excluding procedures marked as not applicable $\left(\frac{Total\ number\ of\ procedures\ observed}{Total\ number\ of\ provedures\ [observed\ and\ not\ observed]} \times 100\right)$.

For each Likert scale item on the teacher survey, the study team calculated the percentage of responses in each category (Agree, Somewhat Agree, Somewhat Disagree, Disagree; Very Useful, Useful, Somewhat Useful, Not Useful).

Qualitative analysis. Content analysis procedures (from Miles et al., 2014) were used to analyze data from the transcripts of the focus groups in order to identify common themes, pitfalls, and benefits. Both a priori and emerging codes were used to identify meaning in the data. Examples of a priori codes are "The GloSS and IKAN helped me modify mathematics interventions," and "The GloSS and IKAN helped me with my day-to-day teaching"; examples of emerging codes are specific types of problems encountered when using the GloSS and IKAN and other feedback regarding use of the GloSS and IKAN. Open coding was used to identify initial codes, followed by an iterative process (axial coding) to search for themes. Identified themes were verified by checking

[1] The Technical Working Group concluded that a sample of 60 students was more than adequate for establishing interassessor reliability as percent agreement (a descriptive statistic).

for their occurrence in multiple sources of data, including the teacher survey responses. Two members of the study team coded the transcripts and discussed any discrepancies until they reached consensus.

In addition, the study team reviewed the data from the optional open-response item on the teacher survey for teacher feedback on the adequacy of training, usefulness of the GloSS and IKAN assessments for guiding interventions, and any problems that arose in administering the assessments. Any teacher feedback was compared with the codes determined in the content analysis of the focus group transcripts.

References

Fannin County School District. (2020). *Fannin County Schools*. Retrieved June 3, 2020, from http://www.fannin.k12.ga.us.

Jefferson County School Board of Education. (2020). *Jefferson County School*. Retrieved June 3, 2020, from https://www.jefferson.k12.ga.us.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). Qualitative data analysis: A methods sourcebook. Sage.

The Governor's Office of Student Achievement. (2020). *Report card, dashboards, & data: Downloadable data*. Retrieved June 3, 2020, from https://gosa.georgia.gov/report-card-dashboards-data/downloadable-data.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2020). *Common Core of Data*. Retrieved June 3, 2020, from https://nces.ed.gov/ccd/schoolsearch/.

Walker County School Systems. (2020, June 3). *Walker Schools*. Retrieved June 3, 2020, from http://www.walkerschools.org.

Appendix C. Supplementary analyses

This appendix includes tables presenting detailed results from the survey and focus groups and from the analysis of the fidelity of administration.

Table C1. Teacher responses to the survey on their perceptions of the usefulness of the Global Strategy Stage (GloSS) and the Individual Knowledge Assessment of Number (IKAN) diagnostic assessments for informing instruction, 2019

Teacher survey questions	Percent of teachers
Based on your experiences using the IKAN and GloSS, together, indicate your level of agreement.	Agree / Somewhat agree / Somewhat disagree / Disagree
1. Data from the IKAN and GloSS are more useful than screening data from assessments used in our school (e.g., STAR, Easy CBM, AIMSweb) for guiding decisions about which students require intervention.	50% / 43% / 7%
2. IKAN and GloSS data help me address struggling students' instructional needs in my day-to-day teaching.	57% / 43%
3. Data from the IKAN are more useful than data from the GloSS.	23% / 20% / 40% / 17%
4. I prefer using mathematics assessment tools **other** than the IKAN and GloSS for guiding instructional decisions.	10% / 43% / 27% / 20%
5. The time that it takes to individually administer the IKAN and GloSS to each student is worth it because I am able to learn more about how each student reasons and thinks about mathematics.	63% / 30% / 7%
Based on your experiences using the GloSS, rate the following items in terms of their usefulness. I find the data on strategy use from the GloSS useful…	Very useful / Useful / Somewhat useful / Not useful
6. …for identifying skills and concepts in which students are weak.	50% / 37% / 13%
7. …for placement within the RtI or multi-tiered system of support (MTSS).	40% / 30% / 27%
8. …for modifying instruction for my mathematics class.	27% / 50% / 20%
9. Overall, I find the data on strategy use from the GloSS useful.	37% / 50% / 13%
Based on your experiences using the GloSS, indicate your level of agreement.	Agree / Somewhat agree / Somewhat disagree / Disagree
10. Administering the GloSS is a good use of my time.	50% / 47%
11. Data on strategy use from the GloSS are useful.	70% / 30%
12. I have enough time to administer the GloSS to all the students in my class.	40% / 17% / 20% / 23%

Based on your experiences using the IKAN, rate the following items in terms of their usefulness. I find the data on strategy use from the IKAN useful…	Very useful / Useful / Somewhat useful / Not useful
13. …for identifying skills and concepts in which students are weak.	47% / 27% / 23%
14. …for placement within the RtI or multi-tiered system of support (MTSS).	40% / 17% / 30% / 13%
15. …for modifying instruction for my mathematics class.	33% / 33% / 23% / 10%
16. Overall, I find the data from the IKAN useful.	40% / 37% / 17% / 7%
Based on your experiences using the IKAN, indicate your level of agreement.	Agree / Somewhat agree / Somewhat disagree / Disagree
17. Administering the IKAN is a good use of my time.	60% / 23% / 10% / 7%
18. Data on knowledge level from IKAN are useful.	50% / 30% / 13% / 7%
19. I have enough time to administer the IKAN to all the students in my class.	57% / 23% / 7% / 13%
Based on your attendance at the IKAN and GloSS training/professional development in February, indicate your level of agreement.	Agree / Somewhat agree / Somewhat disagree / Disagree
20. The training/professional development I received adequately prepared me to properly administer the IKAN and GloSS to my students.	53% / 37% / 7%
21. The training/professional development I received on the IKAN and GloSS was adequate in helping me identify developmental progressions of students' number knowledge and their ability to solve problems involving operations (e.g., addition, subtraction, etc.).	60% / 33% / 7%

STAR, Easy CBM, and AIMSweb are assessments used to identify which students are struggling with mathematics.
Note: *n* = 30. The survey was administered to teachers in the study sample in Fannin County, Jefferson County, and Walker County.
Source: Authors' analysis of primary data collected for the study in 2019.

Table C2. Focus group comments by sample teachers in three counties in Georgia on ways to improve the Global Strategy Stage (GloSS) and Individual Knowledge Assessment of Number (IKAN) diagnostic assessments, 2019 assessment

Ways to improve the assessments
Improve GloSS instructions
Clarify GloSS scoring
Update GloSS vocabulary and pictures
Make GloSS a paper/pencil assessment
Make the IKAN Written Assessment more grade appropriate
Revise IKAN Written Assessment recording sheet
Make IKAN Written Assessment untimed
Make the assessments digitally administered
Link scoring to interventions
Provide more professional development

Note: Teachers in the study sample in Fannin County, Jefferson County, and Walker County in Georgia participated in the focus groups.
Source: Authors' analysis of primary data collected for the study in 2019.

Table C3. Results of the fidelity of administration for the Global Strategy Stage by sample teachers, by item, 2019

Fidelity item	Percent of teachers
The examiner uses the decision rules correctly to decide whether to continue with the next tasks in the Addition and Subtraction section.	67
The examiner uses the decision rules correctly to decide whether to continue with the next tasks in the Multiplication and Division section.	73
The examiner uses the decision rules correctly to decide whether to continue with the next tasks in the Proportions and Ratios section.	83
The examiner gives appropriate wait time as the student answers each task.	100
The examiner probes the student when he or she is non-responsive or when more explanation of the student's strategy is needed.	100
The examiner reads each task to the student and shows him or her the related task card and manipulatives.	100
The examiner always asks the student to explain his or her strategies.	100
The examiner documents student strategies using the student recording sheet.	93
The examiner correctly identifies the domain stage for Addition and Subtraction.	93
The examiner correctly identifies the domain stage for Multiplication and Division.	97
The examiner correctly identifies the domain stage for Ratios and Proportions.	97
The examiner correctly rates the student's overall Global Strategy Stage as the highest stage overall in any domain.	79

Note: The fidelity of administration was analyzed for Fannin County and Jefferson County. Walker County declined to participate.
Source: Authors' analysis of primary data collected for the study in 2019.

Table C4. Results of the fidelity of administration for the Individual Knowledge Assessment of Number (IKAN) Counting Interview by sample teachers, by item, 2019

Fidelity item	Percent of teachers
The examiner has chosen to administer the IKAN Counting Interview to this student because he or she obtained a Stage Score within Stages 0–3 on the GloSS.	94
The examiner administers all three items of the Forward Number Word Sequence section (items 1–3).	100
The examiner records the date of mastery in the record box if all items (1–3) are correct.	100
The examiner administers both items of the Backward Number Word Sequence section (items 4 and 5).	100
The examiner records the date of mastery in the record box if both items (3 and 4) are correct.	89
The examiner administers the number recognition to 1000 portion of the Recognition and Sequence section.	100
The examiner records the date of mastery in the record box if all "What is the Number" items (6–17) are mastered.	80
The examiner administers the "After" number recognition to 1000 of the Recognition and Sequence section.	100
The examiner records the date of mastery in the record box if all "What Number Comes After" items (6–17) are mastered.	89
The examiner administers the "Before" number recognition to 998 of the Recognition and Sequence section.	100
The examiner records the date of mastery in the record box if all "What Number Comes Before" items are mastered.	86
The examiner uses the End of Year Number Knowledge and GSE Expectations matrix and accurately assigns an overall Number Knowledge Stage Score.	89
The examiner annotates student errors as documentation for follow-up instruction.	88
The examiner uses the number cards as prescribed in the examiner's manual.	100

GloSS is Global Strategy Stage. GSE is Georgia Standards of Excellence.
Note: The fidelity of administration was analyzed for Fannin County and Jefferson County. Walker County declined to participate.
Source: Authors' analysis of primary data collected for the study in 2019.

Table C5. Results of the fidelity of administration for the Individual Knowledge Assessment of Number (IKAN) Written Assessment by sample teachers, by item, 2019

IKAN Written Assessment fidelity item	Percent of teachers
The examiner has chosen to administer the IKAN Written Assessment for students who score within Stages 4–8 on the GloSS.	100
The examiner uses the flash or .mp4 video provided.	100
The examiner correctly identifies the overall Number Knowledge Stage Score.	100
The examiner correctly identifies the Stage Score for the Number Sequence and Order Domain.	75
The examiner correctly identifies the Stage Score for the Fractions Domain.	67
The examiner correctly identifies the Stage Score for the Place Value Domain.	75
The examiner correctly identifies the Stage Score for the Basic Facts Domain.	75

GloSS is Global Strategy Stage.
Note: The fidelity of administration was analyzed for Fannin County and Jefferson County. Walker County declined to participate.
Source: Authors' analysis of primary data collected for the study in 2019.