CEMETR-2023-06
SEPT 2023

CEME Technical Report

The Center for Educational Measurement and Evaluation

An Evaluation of the Association Between the Use of the Ignite by Hatch™ Educational Gaming System and the Developmental Status of Young Children Participating in the Georgia DECAL Summer Transition Program

Richard G. Lambert

A PUBLICATION OF THE CENTER FOR EDUCATIONAL MEASUREMENT AND EVALUATION

THE CENTER FOR EDUCATIONAL MEASUREMENT AND EVALUATION

An Evaluation of the Association Between the Use of the Ignite by Hatch™ Educational Gaming System and the Developmental Status of Young Children Participating in the Georgia DECAL Summer Transition Program

Richard G. Lambert, PhD

Center for Educational Measurement and Evaluation

The University of North Carolina at Charlotte

September 2023

An Evaluation of the Association Between Use of the Ignite by Hatch™ Educational Gaming System and the Developmental Status of Young Children Participating in the 2023 Georgia DECAL Summer Transition Program

This study examined the relationship between using the Ignite by Hatch™ educational gaming system and assessments of the developmental status of young children participating in the 2023 Georgia Department of Early Care and Learning (DECAL) Summer Transition Program. Researchers collected data that describe the extent to which the children engaged with Ignite system and the skill levels they achieved within the system. Teachers also rated the developmental status of the children using rating scales that focused on Literacy and Mathematics skills.

Assessment Measures

Researchers from Hatch™ created two new pilot assessment measures for this project. The two measures, one focused on Literacy skills and the other on Mathematics skills, consist of developmental progressions presented as rating scales for teachers. Researchers designed the progressions to help teachers rate the developmental status of young children (ages 4 to 5 years old). Teachers who use the assessment measures rate children based on evidence and observational information gathered in the course of instruction. Each progression is based on a specific instructional objective from the Georgia Early Learning and Development Standards (GELDS) for young children and contains specific behavioral anchors for each of four developmental levels: "Not Yet," "Beginning," "Developing," and "Demonstrating." Children rated by their teacher at the "Demonstrating" level meet or exceed the expectations outlined by each of the applicable GELDS. Appendix B of this report contains the instructional objectives from the GELDS that align with each developmental progression along with the full text of each progression. Given that these measures are new, we will first focus on an evaluation of the psychometric properties of the information that the measures provided.

Results From the Psychometric Analysis of the Literacy Assessment

Dimensionality

All Rasch modeling analyses were conducted using the Winsteps software platform (Linacre, 2020). Rasch modeling assumes what is called *unidimensionality*. In this case, this term implies that the developmental progressions that make up the Literacy assessment measure one and only one underlying latent construct. Unidimensionality was evaluated using exploratory factor analysis through the principal axis factoring procedure. One factor with an eigenvalue of 5.976 accounted for 59.76% of the variance in the ratings. No other eigenvalue was greater than 1.0. The scree plot clearly indicated a single-factor solution. The factor loadings for the Literacy assessment progressions ranged from 0.610 to 0.836. The unidimensionality of the Literacy total scale score was also evaluated by using mean-square (MNSQ) progression fit statistics and Rasch principal components analysis of residuals (PCAR). MNSQ fit values between 0.6 and 1.4 are considered reasonable for rating scale progressions (Bond & Fox, 2007). MNSQ values less than 2.0 can indicate that a progression, though not fitting optimally with the measurement model, can still contribute useful information to the overall score on the measure. Infit statistics indicate the fit of individual progression response patterns to the measurement model. They also address the possibility of secondary dimensions and fit to the underlying construct. Outfit statistics are sensitive to outliers—that is, responses that show great differences between individual responses and progression difficulties. They are also sensitive to unusual and unexpected progression response patterns.

For PCAR, a variance of greater than 50% explained by measures is considered good and offers support for scale unidimensionality. If a secondary dimension has an eigenvalue of smaller than 3 and accounts for less than approximately 5% of the unexplained variance, unidimensionality is considered plausible (Linacre, 2012).

The PCAR showed that the Rasch dimension explained most of the variance in the data (57.8%) with an eigenvalue of 13.607, relative to the total eigenvalue of 23.607. The first contrast (the largest potential secondary dimension) had an eigenvalue of 1.60 and accounted for 6.8% of the unexplained variance. The fit statistics for all of the progressions were well within acceptable limits. The infit MNSQ values ranged from 0.80 to 1.58. The outfit MNSQ values ranged from 0.80 to 1.87. The progression to total score correlations, with each progression included in the total score, ranged from .66 to .81. In summary, these model fit statistics, when taken together, generally suggest that the data does in fact fit the Rasch partial-credit model (PCM) very well. These results also indicate that the data satisfied the unidimensionality assumption of the Rasch model.

Item Difficulty Measures

The progression location hierarchy generally appeared to be consistent with the expected developmental trajectory for typically developing prekindergarten children. The progressions pertaining to a child's ability to recognize letters of the alphabet (Q7) and copy print (Q10) were estimated as the relatively easiest progressions (-0.57 and -0.56, respectively). The results indicated that the progression pertaining to a child's ability to segment phonemes (Q5) was the most difficult (0.53). The results indicated that the remaining progressions were average in difficulty level.

The range of progression difficulties (-0.57 to 0.53) was found to be relatively narrow, and it will be ideal to add progressions with a wider range of difficulty levels in the future. However, when the progression rating scale anchor point, or category, locations are considered, these values come very close to matching the range of abilities of the children assessed. These values indicate the ability locations that form the model-estimated boundaries between the rating scale or progression categories. These locations indicate where on the underlying ability scale, or total score, the probability becomes higher that a child will be placed at the next-highest category on the progression relative to the previous anchor point. The values ranged from as low as 1.56 for the boundary between 1 and 2 for Q10 (copy print) to as high as 1.41 for the boundary between 3 and 4 for Q5 (segments phonemes). These values much more closely match the full range of ability estimates on the total score and provide reasonable separation of children according to underlying ability.

In summary, the developmental pathway from the easiest to the most difficult progressions for the Literacy assessment generally aligns with expectations from developmental theory. It is also important to recognize, as indicated, that the range of progression difficulties is effectively much wider than the results indicate when considering the separation created between children by the range of rating scale anchor point threshold locations.

Reliability

Reliability was evaluated using the following Rasch indexes: the person separation index, item separation index, person reliability, and item reliability. Item (progression) and person reliabilities were evaluated using both sample-based and model-based coefficients. The person separation index, an estimate of the adjusted person standard deviation divided by the average measurement error, indicates how well the instrument can discriminate persons on each of the constructs. The item (progression) separation index indicates an estimate in standard error units of the spread or separation of progressions along the measurement constructs. Reliability separation indexes greater than 2 are considered adequate, and indexes greater than 3 are considered high (Bond & Fox, 2007). High person or item (progression) reliability means that there is a high probability of replicating the same separation of persons or progressions across measurements. Specifically, person separation reliability estimates the replicability of person placement across other progressions measuring the same construct. Similarly, progression separation reliability estimates the replicability of progression placement along the construct developmental pathway if the same progressions were given to another sample with similar ability levels. The person reliability provided is similar to the classical or traditional test reliability, whereas the progression reliability has no classical equivalent. Low values in person and progression reliability may indicate a narrow range of person or progression measures. It may also indicate that the number of progressions or the sample size under study is too small for stable estimates (Linacre, 2012). Reliability was also evaluated using Cronbach's alpha measure of internal consistency.

The item (progression) reliability values, both sample-based and model-based, were above 0.99. The item (progression) separation indexes were also very high: sample-based was 9.76 and model-based was 10.28. Taken together, these findings indicate that it is reasonable to expect highly consistent estimates of progression difficulty levels across samples. The sample-based person separation index was 2.32 and the model-based value was 2.60. The sample-based person reliability index was 0.84 and the model-based value was 0.87. The Cronbach's alpha value for the total score was .94. Based on these reliability indexes, the total scores appear to yield adequately reliable information from this sample. Specifically, these results indicate that it is reasonable to expect reliable estimates of child overall ability levels when teachers use the progressions to make valid placements for young children and those placements are transformed into a composite or total score. It is important to note that these results address reliability issues related to the use of a total score and may be very different from the results of an inter-rater reliability study.

Rating Scale Category Effectiveness

A rating scale with demonstrated category effectiveness yields evidence that raters are using the scale as it was intended to be used. This means that raters can use the scale to discriminate between responses with true underlying differences on the construct being measured. In this case, rating scale category effectiveness is a measure of the validity of the data elicited by the developmental progressions. Developmental progressions with effective rating scales yield valid data that can be used to place children along a continuum of development so that the placements both reflect the true developmental status of each child and can be used by teachers to differentiate instruction and support growth, learning, and development. Therefore, rating scale category effectiveness was examined to provide information about the rating scale categories on specific progressions and to evaluate whether teachers appear to be using the progressions in the manner intended. Rating scale effectiveness was also examined to evaluate if it is reasonable to apply Rasch modeling to the data.

Not all rating scales are created equal, and not all raters use rating scales effectively. Ideally, rating scale data is most valid when the intended meaning of each of the individual rating scale anchor points is communicated clearly and unambiguously to respondents or raters and when raters use the scales as intended. The evaluation of rating scale category effectiveness can suggest the optimal number of rating scale categories, places along the scale where categories can be combined, and categories that may be misunderstood or misused by raters. In this case, valid placements can only occur when teachers understand the purpose of the assessment, are well trained, understand the intended content of both the progressions and their rating scale anchor points, and collect and analyze valid evidence and observations to support placements on the progressions. An examination of rating scale effectiveness can help identify potential problems with the progressions or their use. However, further research is often needed to determine whether identified problems are related to the progressions themselves, their use by raters, the quality of rater training, or some combination of these factors.

It is recommended that for each progression, each rating scale category is assigned to a minimum of 10 children. All rating scale categories should be used by the raters, and each rating scale category should be assigned to enough children to allow for reasonable statistical estimates within the Rasch modeling process. These criteria were easily met for all 10 Literacy progressions.

Appendix A contains a table that shows the percentage of children assigned to each level for each progression. Every rating scale category was used at least 100 times for each progression.

Next, the overall ability estimates, based on the total scale scores, were examined for all children in the sample who were placed at a particular response category or scale point on each of the developmental progressions. Average measure scores should advance monotonically with rating scale category values (Bond & Fox, 2007). This criterion was met for all the Literacy progressions. Essentially, this finding indicates that children who were rated as "Demonstrating" on a particular progression scored higher, on average, on the total Literacy score than children who were rated as "Developing." Similarly, the average total score for children who were rated as "Developing" was higher than those who were rated as "Beginning," who, in turn, scored higher than those who were rated as "Not Yet." This finding is a very strong indicator that teachers were generally using the rating scale correctly.

Next, the category thresholds were examined. Thresholds (also called step calibrations) are the difficulty levels estimated at the point on the total score at which teachers are more likely to choose one response category or rating scale point over the previous step on the progression (Bond & Fox, 2007). For this study, the Andrich thresholds from the Rasch PCM were used. Thresholds should also increase monotonically along the rating scale categories. The model-based thresholds advanced as expected for all Literacy progressions, and this, too, is a positive finding, indicating some evidence that teachers were using the rating scale correctly.

All of these favorable findings taken together supported the development of Literacy standard scores based on the raw scores from the developmental progressions. The ratings data from each progression consisted of ordinal ratings. The Rasch PCM was used to convert the sum of the raw ordinal ratings across the progressions into a composite total score. This total score, or standard score, spread the children out across an interval scale with a mean of 500 and a standard deviation of 100.

Results From the Psychometric Analysis of the Mathematics Assessment

Dimensionality

Unidimensionality was evaluated using exploratory factor analysis through the principal axis factoring procedure. One factor with an eigenvalue of 5.819 accounted for 58.19% of the variance in the ratings. No other eigenvalue was greater than 1.0. The scree plot clearly indicated a single-factor solution. The factor loadings for the Literacy assessment progressions ranged from 0.711 to 0.840. The PCAR showed that the Rasch dimension explained most of the variance in the data (60.9%) with an eigenvalue of 15.557, relative to the total eigenvalue of 25.557. The first contrast (the largest potential secondary dimension) had an eigenvalue of 1.59 and accounted for 6.2% of the unexplained variance. The fit statistics for all of the progressions were well within acceptable limits. The infit MNSQ values ranged from 0.77 to 1.20. The outfit MNSQ values ranged from 0.72 to 1.25. The progression to total score correlations, with each progression included in the total score, ranged from .67 to .83. In summary, these model fit statistics, when taken together, generally suggest that the data does in fact fit the Rasch PCM very well. These results also indicated that the data satisfied the unidimensionality assumption of the Rasch model.

Item Difficulty Measures

The progression location hierarchy appeared to be generally consistent with the expected developmental trajectory for typically developing prekindergarten children. The progressions pertaining to a child’s ability to sort (Q8), count (Q2), and recite numbers (Q6) were estimated as the relatively easiest progressions (-0.81, -0.78, and -0.76, respectively). The results indicated that the progressions pertaining to a child's ability to solve mathematical problems (Q10) and use graphs

(Q4) were considered relatively difficult (1.79 and 0.99, respectively). The results indicated that the remaining progressions were average in difficulty level.

The range of progression difficulties (-0.81 to 1.79) represented a reasonably wide range of difficulty. When the progression rating scale anchor point, or category, locations are considered, these values come very close to matching the range of abilities of the children assessed. The values ranged from as low as -1.81 for the boundary between 1 and 2 for Q9 (directional language) to as high as 2.20 for the boundary between 3 and 4 for Q10 (problem-solving). These values closely match the full range of ability estimates on the total score and provide reasonable separation of children according to underlying ability.

In summary, the developmental pathway from the easiest to the most difficult progressions for the Mathematics assessment generally aligns with expectations from developmental theory. It is also important to recognize, as indicated, that the range of progression difficulties is effectively much wider than the results indicate when considering the separation created between children by the range of rating scale anchor point threshold locations.

Reliability

The item (progression) reliability values, both sample-based and model-based, were above 0.99. The item (progression) separation indexes were also very high: sample-based was 9.76 and model-based was 10.28. Taken together, these findings indicate that it is reasonable to expect highly consistent estimates of progression difficulty levels across samples. The sample-based person separation index was 2.32 and the model-based value was 2.60. The sample-based person reliability index was 0.84 and the model-based value was 0.87. The Cronbach's alpha value for the total score was .94. Based on these reliability indexes, the total scores appear to yield adequately reliable information from this sample. Specifically, these results indicate that it is reasonable to expect reliable estimates of child overall ability levels when teachers use the progressions to make valid placements for young children and those placements are transformed into a composite or total score. It is important to note that these results address reliability issues related to the use of a total score and may be very different from the results of an inter-rater reliability study.

Rating Scale Category Effectiveness

It is recommended that for each progression, each rating scale category is assigned to a minimum of 10 children. All rating scale categories should be used by the raters, and each rating scale category should be assigned to enough children to allow for reasonable statistical estimates within the Rasch modeling process. These criteria were easily met for all 10 Mathematics progressions. Appendix A contains a table that shows the percentage of children assigned to each level for each progression. Every rating scale category was used at least 100 times for each progression.

Next, the overall ability estimates, based on the total scale scores, were examined for all children in the sample who were placed at a particular response category or scale point on each of the developmental progressions. Average measure scores should advance monotonically with rating scale category values (Bond & Fox, 2007). This criterion was met for all the Mathematics progressions. As explained for the Literacy progressions, this finding for the Mathematics progressions indicates that children who were rated as "Demonstrating" on a particular progression scored higher, on average, on the total Mathematics score than children who were rated as "Developing." Similarly, the average total score for children who were rated as "Developing" was higher than those who were rated as "Beginning," who, in turn, scored higher than those who were rated as "Not Yet." This finding is a very strong indicator that teachers were generally using the rating scale correctly. In addition, the model-based thresholds advanced as expected for all Mathematics progressions, and this, too, is a very positive finding, indicating some evidence that teachers were using the rating scale correctly.

All of these favorable findings taken together supported the development of Mathematics standard scores based on the raw scores from the developmental progressions. The ratings data from each progression consisted of ordinal ratings. The Rasch PCM was used to convert the sum of the raw ordinal ratings across the progressions into a composite total score. This total score, or standard score, spread the children out across an interval scale with a mean of 500 and a standard deviation of 100.

Summary of Psychometric Analyses

This initial investigation of the psychometric properties of the information provided by the Literacy and Mathematics assessment measures resulted in many favorable indications that the measures have promise and may be useful tools for teachers and researchers. These results support the creation of Literacy and Mathematics interval-level standard scores. The initial reliability and validity statistics offer evidence that supports the usefulness of the information the measures can provide as a source of indicators of child developmental status from the perspective of teachers. It is important to note that this study did not include an examination of the inter-rater reliability or the strictness, leniency, or bias of the raters. Future research is needed that can accumulate more evidence to support specific uses of these measures for research or instructional planning purposes.

Initial Descriptive Examination of the Correlations Between Study Variables

Table 1 contains the correlation coefficients that describe the relationships between time spent engaged with the Ignite system, skill levels achieved within the gaming system, and the assessment scores. The total time a child spent engaging with the Ignite educational gaming system was strongly, positively correlated with time spent playing both Literacy games (.942) and Mathematics games (.936). Therefore, we can conclude that if children in this study spent time engaged with the Ignite system, they played Literacy and Mathematics games. Furthermore, children who spent more time playing Literacy and Mathematics games tended to achieve higher skill levels within the gaming system. For example, time spent playing Literacy games was strongly correlated with the Literacy level achieved (.746). Time spent playing Mathematics games was moderately correlated with the Mathematics level achieved (.543). These findings illustrate how many children advanced through the levels within the gaming system as they spent more time engaged with the system. However, playtime alone is not a perfect predictor of level achieved. Some children spent time in the system without progressing through the levels.

Playtime alone was not strongly related to the assessment scores. Specifically, time spent playing Literacy games was weakly correlated with Literacy standard scores (.140). Time spent playing Mathematics games was not related to Mathematics standard scores (.055). Literacy level achieved was weakly correlated with Literacy standard scores (.126). Mathematics level achieved was also weakly correlated with Mathematics standard scores (.151).

Relationship Between Minutes Engaged With Ignite and Levels Achieved for All Children

A total of 3,736 children in the Georgia DECAL Summer Transition Program engaged with the Ignite system. Table 2 contains a more detailed description of the strong relationship between minutes engaged with the Ignite system and Literacy skill levels achieved for all children who attended the summer program and engaged with the Ignite system. Table 3 contains similar information for the moderately strong relationship between time spent playing Mathematics games and the Mathematics levels achieved. In the Language & Communication domain, children who achieved only Level 1 played Language games for an average of 7.34 minutes. In contrast, children who achieved Level 4 played Language games for an average of 31.16 minutes. In the Literacy domain, children who achieved only Level 1 played Literacy games for an average of 6.92 minutes. In contrast, children who achieved Level 4 played Literacy games for an average of 90.93 minutes. In the Mathematics domain, children who achieved only Level 1 played Mathematics games for an average of 19.98 minutes. In contrast, children who achieved Level 4 played Mathematics games for an average of 66.97 minutes. These results confirm some basic patterns of usage. Playtime alone does not automatically lead to higher levels achieved and skills mastered. However, when children engage with the Ignite system in a meaningful way for more time, they tend to achieve higher levels within the system and gain mastery of more skills.

Not all children who attended the Georgia DECAL Summer Transition Program and engaged with the Ignite system (N = 3,736) were assessed by their teachers using the pilot assessment measures. Teachers assessed, using either or both the Literacy and Mathematics assessment measures, only 45.3% of the program's children who engaged with the Ignite Literacy games. Similarly, teachers assessed, using either or both the Literacy or Mathematics assessment measures, only 43.4% of the program's children who engaged with the Ignite Mathematics games. Therefore, we examined only those children who were assessed by their teacher to determine if the same patterns illustrated by Tables 2 and 3 were also sustained. It was important to determine if there were systematic differences between children who were and were not assessed by their teacher.

Table 4 contains a more detailed description of the strong relationship between minutes engaged with the Ignite system and Literacy skill levels achieved for all children who were assessed. Table 5 contains similar information for the moderately strong relationship between time spent playing Mathematics games and Mathematics levels achieved for all children who were assessed by their teacher. In the Language & Communication domain, children who achieved only Level 1 played Language games for an average of 8.30 minutes. In contrast, children who achieved Level 4 played Language games for an average of 32.78 minutes. In the Literacy domain, children who achieved only Level 1 played Literacy games for an average of 7.77 minutes. In contrast, children who achieved Level 4 played Literacy games for an average of 96.27 minutes. In the Mathematics domain, children who achieved only Level 1 played Mathematics games for an average of 21.71

minutes. In contrast, children who achieved Level 4 played Mathematics games for an average of 69.91 minutes. These results confirm that the same basic patterns of usage observed for all children in the Georgia DECAL SummerTransition Program were observed when focusing only on children who were assessed by their teacher.

Next, we examined whether children who attended classrooms where the teacher completed the assessments engaged with the Ignite system more frequently. Children who were assessed with the Literacy assessment engaged with the Ignite system for an average of 131.74 minutes (SD = 74.17). Children who were not assessed with the Literacy assessment engaged with the games for an average of 117.53 minutes (SD = 72.08). This small difference was statistically significant (t = 5.848, p < .001). Similarly, children who were assessed with the Mathematics assessment engaged with the Ignite system for an average of 132.20 minutes (SD = 73.72). Children who were not assessed with the Mathematics assessment engaged with the games for an average of 117.42 minutes (SD = 72.38). This small difference was also statistically significant (t = 6.072, p < .001). These differences were very similar because most children who were assessed were assessed using both the Literacy and Mathematics measures. It is possible that teachers who completed the assessments were somewhat more likely to emphasize the Ignite system in their classrooms.

Relationship Between Levels Achieved and Assessment Scores

We examined the relationship between the levels that children achieved within the Ignite system and the standard scores they received on both the Literacy and Mathematics assessments. We conducted these analyses with several more nuanced strategies than simply examining the correlation coefficients, as described earlier in this report. First, we identified the maximum level a child achieved across all the games in the Literacy domain and again in the Mathematics domain. Next, children who achieved Level 4 or higher were combined into one group for each domain. There were two reasons for this decision: Level 4 represents kindergarten readiness and very few numbers of children achieved Levels 5, 6, 7, or 8. This process identified a strong functional relationship between levels and standard scores for both Literacy and Mathematics.

There was a gain in standard scores, on average, for every additional level achieved (see Table 6). Specifically, children who achieved a maximum of Level 1 in the Literacy domain scored, on average, 479.03 (*SD* = 112.38) on the Literacy assessment. Children who achieved a maximum of Level 2 in the Literacy domain scored, on average, 502.77 (*SD* = 97.80) on the Literacy assessment. Children who achieved a maximum of Level 3 in the Literacy domain scored, on average, 523.83 (*SD* = 82.08) on the Literacy assessment. Children who achieved a maximum of Level 4 or higher in the Literacy domain scored, on average, 536.83 (*SD* = 69.91) on the Literacy assessment.

The data revealed a similar pattern for Mathematics. Children who achieved a maximum of Level 1 in the Mathematics domain scored, on average, 479.73 (*SD* = 103.47) on the Mathematics assessment. Children who achieved a maximum of Level 2 in the Mathematics domain scored, on average, 495.63 (*SD* = 105.20) on the Mathematics assessment. Children who achieved a maximum of Level 3 in the Mathematics domain scored, on average, 511.55 (*SD* = 94.02) on the Mathematics assessment. Children who achieved a maximum of Level 4 or higher in the Mathematics domain scored, on average, 528.41 (*SD* = 80.05) on the Mathematics assessment.

These results indicate a moderately strong association between the maximum level achieved and standard scores for both the Literacy and Mathematics teacher ratings of developmental status. Next, we considered children who engaged with the Ignite system but only achieved Level 1 as the comparison condition. These children engaged with a computer-based educational experience but did not master any age-appropriate skills. We considered children who engaged with the games and achieved Level 4 or higher as the treatment condition. These children engaged with the computer-based educational experience and mastered age-appropriate skills. When these two groups were compared, we observed moderately large effect sizes in favor of the treatment-condition children.

Children who engaged with the games and achieved Level 4 or higher in Literacy scored approximately one-half standard deviation higher on the Literacy assessment (effect size = .550) than children who remained at Level 1. Children who achieved Level 4 or higher in Mathematics also scored approximately one-half standard deviation higher on the Mathematics assessment (effect size = .505) than children who remained at Level 1. These differences were statistically significant ($p < .001$).

A final analysis was completed to examine the relationship between the level achieved and the assessment scores. When all children who achieved a specific level or range of levels are included in the same group, as shown in Table 6, the analyses ignore potential rater or site effects. Therefore, for the final analysis, children were nested within their site in the context of multilevel modeling. Level 1 in the model consisted of the maximum level a child achieved as the predictor variable and the standard score as the outcome measure. Level 2 consisted of the site where the children attended the Georgia DECAL Summer Transition Program. Demographic covariate or predictor variables were not available for the children and therefore could not be included in the Level 1 models. Similarly, features of the sites, classrooms, or teachers were not available as covariates or predictors for the Level 2 models. Almost all the children attended the program at sites throughout Georgia that housed only one classroom. Therefore, Level 2 represents the site, teacher, or rater. In a few cases, there were multiple classrooms within a site. A total of 107 sites contributed assessment data for 1,689 children (average = 15.9 children per site) to this study. We treated children who did not progress beyond Level 1 as the baseline or comparison condition. We treated children who achieved Levels 2, 3, and 4 or higher as dummy-coded comparison groups. Each level achieved was treated as a group-mean-centered, dummy-coded, fixed-effects predictor variable in the models.

First, we examined the unconditional models (no predictor variables). The results of the unconditional multilevel models indicated that the majority of the variance in the standard scores for both Literacy (57.7%) and Mathematics (56.7%) was between children within the same site. However, a substantial portion of the variance in both standard scores was between sites for both Literacy (42.3%) and Mathematics (43.3%). It is possible that these between-site differences were associated with differences in the demographic background of the children and families served, community resources, instructional strategies, or rate effects at each site. Future research will need to be conducted that includes a range of variables that represent site features to further understand this finding. However, it is important to note that the presence of substantial between-site differences justifies the use of multilevel modeling and indicates that the whole group analyses contained within Table 6 may not reveal a complete picture of the results.

Table 7 displays the results of the multilevel conditional models. These results display the pooled effects across sites. First, we entered total time spent engaged with the Ignite games as the predictor of the standard scores. There was a statistically significant positive association between time spent playing the games and the standard scores for both Literacy ($p < .01$; pseudo $r^2 = .018$, pseudo $r = .136$) and Mathematics ($p < .01$; pseudo $r^2 = .019$, pseudo $r = .139$). Next, we entered the dummy variables that represent levels achieved above the baseline level, Level 1. There were statistically significant effects, relative to those children who remained at Level 1, for Levels 2, 3, and 4 or higher for both Literacy and Mathematics.

The model predicted, when considering site differences, that children who achieved Level 2 in the Literacy domain would score 41.23 points higher than Level 1 children ($p < .001$), Level 3 children would score 77.03 points higher ($p < .001$), and children who reach a Level 4 or higher would score 92.66 points higher ($p < .001$). With the addition of the level achieved in the Literacy model, time spent was no longer significantly associated with Literacy standard scores, and the coefficient became negative (-.01). Furthermore, the magnitude of the effects for the level achieved was much greater (pseudo $r^2 = .051$, pseudo $r = .226$) than that observed with the descriptive statistics alone (see Table 6). Similarly, the Literacy standard-score difference between Level 1 and Level 4 was close to a full SD unit. This finding demonstrates that when the time spent engaged with the games is controlled for, the level achieved has an even stronger association with the Literacy standard scores.

Similarly, for the Mathematics standard scores, the model predicted, when considering site differences, that children who achieved Level 2 in the Mathematics domain would score 18.92 points higher than Level 1 children ($p < .01$), Level 3 children would score 53.50 points higher ($p < .001$), and children who reached Level 4 or higher would score 97.47 points higher ($p < .001$). With the addition of the level achieved in the Mathematics model, time spent was still statistically significant ($p < .05$), but the coefficient became negative (-.09). Furthermore, as seen with the Literacy results, the magnitude of the effects for the level achieved was much greater (pseudo $r^2 = .137$, pseudo $r = .370$) than that observed with the descriptive statistics alone (see Table 6). Similarly, the Mathematics standard-score difference between Level 1 and Level 4 was close to a full SD unit. This finding demonstrates again that when the time spent engaged with the games is controlled for, the level achieved has an even stronger association with the Mathematics standard scores.

We tested whether these effects for the level achieved could be treated as random effects. This test indicates whether the predicted advantages for levels achieved were consistent or variable across sites. There was not sufficient between-site variance in the magnitude of the effects to justify a random-effects approach to these predictor variables. This finding supports the interpretation that these effects were consistent across sites. Figure 1 shows the difference between Level 1 and Level 4 children in terms of their Literacy standard scores. Each line in the graph illustrates the slope for the Level 4 effect for a specific site. The overall pattern in the graph illustrates that these slopes' values were consistent across sites and independent of whether a teacher was strict or lenient with their ratings of child developmental status or whether there are between-site differences with respect to the initial developmental status of the children. It is also important to note that the magnitude of the model-estimated effects for each level is somewhat larger than that reported in Table 6. This finding suggests that including the between-site differences and time spent playing games in the model made a meaningful difference in the strength of the findings.

Figure 2 illustrates the same pattern for the Mathematics standard scores. The slope of the Level 4 effect was consistent across sites, and there was not sufficient variance in the slopes to justify a random-slopes model. Therefore, these results demonstrate that the magnitude of the advantage in Mathematics standard scores for the children who reached Level 4 was consistent whether the teacher was strict or lenient in their ratings or whether the children had relatively lower or higher developmental levels.

Conclusion

The results of this study demonstrated that children in the Georgia DECAL Summer Transition Program generally could be expected to reach age-appropriate levels of skill within the Ignite system when they engage with the system for enough time. This study also demonstrated that time in the system is not enough to ensure that children will make progress; children also have to engage with the games in meaningful ways. This study also demonstrated that children who do advance within the Ignite system to age-appropriate levels (Level 4 or higher) tend to have higher scores in Literacy and Mathematics skill development, as measured by the pilot measures used in this study. Therefore, the results of this study suggest that the pathway from time spent engaged within the Ignite system to higher assessment scores is mediated by the extent to which children achieve higher skill levels within the Ignite system.

There are some important limitations to the findings of this study. First, the assessment measures are new, and more research is needed to continue to accumulate evidence for the usefulness of the information they provide for both research and instructional purposes. Second, it is important to note that this was an observational study, and there was no random assignment of children to treatment and control conditions. There may be various preexisting differences between children who reached higher levels (Level 4 or higher) in the Ignite system and those children who remained at Level 1. This study suggests the need to examine in future research whether children of higher initial skills are more likely to achieve higher levels within the gaming system. While this study explored and demonstrated some meaningful associations between the levels achieved within the Ignite system and the assessment scores, it remains unclear how children would have progressed without any exposure to Ignite. Future research will need to randomly assign children to receive or not receive Ignite experiences so that clearer attributions can be made that can connect developmental status to engagement with the Ignite system.

Third, when we consider that not all teachers completed the assessments, it is important to recognize that these results may not generalize to all children. It is possible that teachers who completed the assessments emphasized the use of the Ignite system in their classrooms more than teachers who did not complete the assessments. It is also possible that teachers who emphasize assessment and Ignite in their teaching can be more lenient or stricter in their assessment practices. Future research is needed to examine the inter-rater reliability of teachers as sources of evidence about child developmental status using measures like those employed in this study.

References

Bond, T. G. & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the Human sciences* (2nd ed.). Lawrence Erlbaum.

Linacre, J. M. (2020). *Winsteps* (Version 4.6.2) [Computer software]. Winsteps.

Linacre, J. M. (2012). *A user's guide to Winsteps*. Winsteps. https://www.winsteps.com/winman/copyright.htm

Table 1

Correlations Between Study Variables

	Total PT	Mathematics PT	Literacy PT	Language PT	Literacy SS	Mathematics SS	Literacy level
Mathematics PT	.936**						
Literacy PT	.942**	.824**					
Language & Communication PT	.815**	.726**	.731**				
Literacy SS	.095**	.070**	.140**	-0.013			
Mathematics SS	.089**	.055*	.145**	-0.007	.889**		
Literacy level	.698**	.597**	.746**	.528**	.126**	.124**	
Mathematics level	.576**	.543**	.643**	.475**	.133**	.151**	.682**

Note. PT = playtime; SS = standard score; * = $p < .05$; ** = $p < .01$.

Table 2
Playtime by Maximum Literacy Level Achieved - All Children

Domain	Maximum Literacy level achieved	*n*	Playtime Mean	Playtime *SD*
All domains	1	705	45.18	27.56
	2	2,629	128.83	48.74
	3	282	192.07	51.66
	4+	120	322.19	125.83
	Total	3,736	124.03	72.95
Language & Communication	1	705	7.34	4.33
	2	2,629	14.19	6.01
	3	282	18.47	7.03
	4+	120	31.16	14.79
	Total	3,736	13.76	7.68
Literacy	1	705	6.92	3.93
	2	2,629	27.41	11.78
	3	282	51.25	12.52
	4+	120	90.93	39.02
	Total	3,736	27.38	20.31

Table 3

Playtime by Maximum Mathematics Level Achieved - All Children

Domain	Maximum Mathematics level achieved	*n*	Playtime Mean	Playtime *SD*
All domains	1	785	74.67	38.72
	2	893	115.41	38.22
	3	1,408	148.02	51.13
	4+	374	229.34	103.77
	Total	3,460	131.75	69.91
Mathematics	1	785	19.98	16.50
	2	893	41.49	17.25
	3	1,408	52.46	20.96
	4+	374	66.97	32.83
	Total	3,460	43.83	25.54

Table 4

Playtime by Maximum Literacy Level Achieved - Assessed Children

Domain	Maximum Literacy level achieved	*n*	Playtime Mean	*SD*
All domains	1	232	52.72	29.54
	2	1,274	131.03	45.97
	3	127	188.69	48.20
	4+	61	345.18	143.90
	Total	1,694	132.34	73.18
Language & Communication	1	232	8.30	5.00
	2	1,274	14.38	5.77
	3	127	17.85	6.56
	4+	61	32.78	15.95
	Total	1,694	14.47	7.67
Literacy	1	232	7.77	4.01
	2	1,274	27.98	11.28
	3	127	51.17	12.15
	4+	61	96.27	46.41
	Total	1,694	29.41	21.15

Table 5

Playtime by Maximum Mathematics Level Achieved - Assessed Children

Domain	Maximum Mathematics level achieved	*n*	Playtime Mean	Playtime *SD*
All domains	1	326	80.04	37.77
	2	415	118.61	36.97
	3	703	147.44	48.49
	4+	178	238.58	119.80
	Total	1,622	136.52	71.69
Mathematics	1	326	21.71	16.29
	2	415	43.14	17.24
	3	703	52.24	19.73
	4+	178	69.91	37.99
	Total	1,622	45.72	25.61

Table 6
Literacy and Mathematics Standard Scores by Level

Maximum Literacy level achieved	*n*	Literacy standard scores Mean	*SD*
1	218	479.03	112.38
2	1,153	502.77	97.80
3	117	523.83	82.08
4+	59	536.83	69.91
Total	1,547	502.32	98.72

Maximum Mathematics level achieved	*n*	Mathematics standard scores Mean	*SD*
1	293	479.73	103.47
2	381	495.63	105.20
3	628	511.55	94.02
4+	151	528.41	80.05
Total	1,453	502.71	98.75

Table 7

Multilevel Modeling Results - Literacy and Mathematics Standard Scores

Domain	Effect	Coefficient	*SE*	*t*	
Literacy	Intercept	500.69	6.70	74.74	***
	Level 2	40.62	9.67	4.20	***
	Level 3	75.70	13.23	5.72	***
	Level 4+	89.92	16.93	5.31	***
Mathematics	Intercept	502.00	7.01	71.64	***
	Level 2	16.32	6.71	2.43	*
	Level 3	47.97	7.78	6.17	***
	Level 4+	83.64	11.10	7.53	***

Note. * - $p < .05$; *** - $p < .001$.

Figure 1

Contrast Between Literacy Standard Scores for Levels 1 and 4 by Site

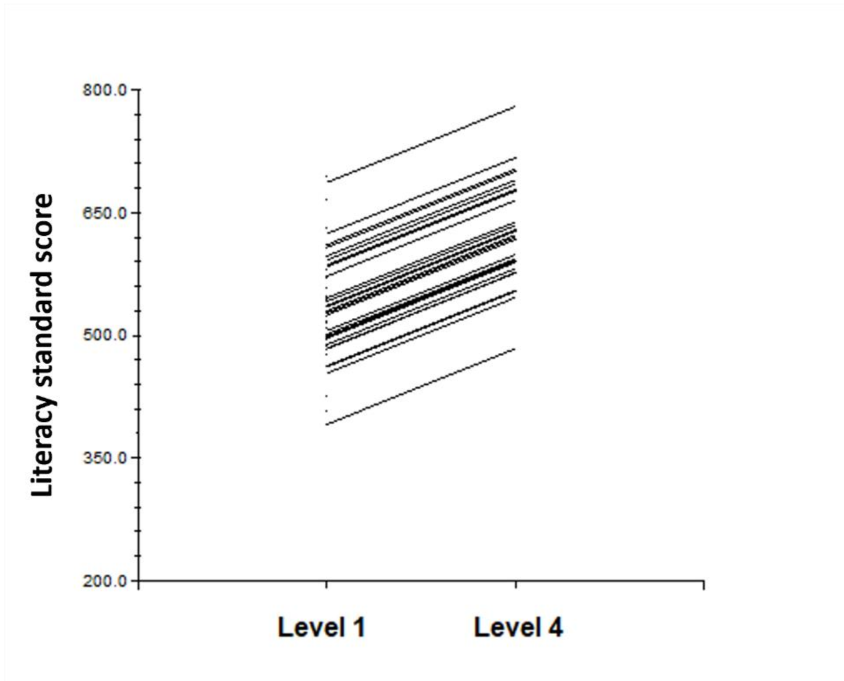
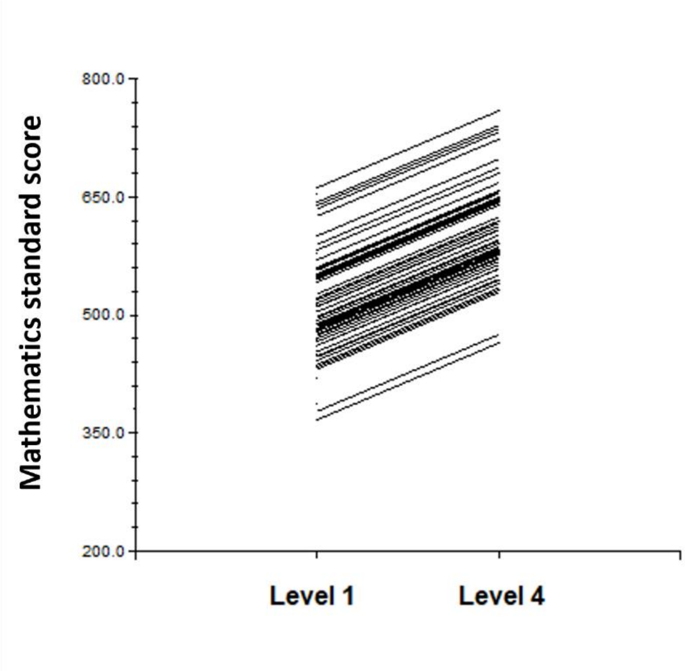


Figure 2

Contrast Between Mathematics Standard Scores for Levels 1 and 4 by Site



Appendix A

Distribution of Teacher Ratings by Progression

Literacy progression	Not Yet	Beginning	Developing	Demonstrating
Predictions	16.0	18.1	25.3	40.6
Connections	9.4	14.6	20.4	55.6
Alternate ending	14.9	17.7	30.1	37.3
Differentiates	12.4	16.2	29.1	42.2
Segments	15.6	24.2	25.7	34.5
Blends	14.9	20.2	26.5	38.4
Alphabet recognition	8.1	14.3	23.9	53.6
Left to right	16.1	19.0	22.8	42.1
Environmental print	10.6	14.7	20.0	54.7
Copies text	7.8	16.9	20.0	55.3

Mathematics progression	Not Yet	Beginning	Developing	Demonstrating
Combines shapes	11.1	13.5	24.4	51.0
Counts	5.8	9.5	17.5	67.1
Patterns	7.3	11.0	22.3	59.4
Graphs	16.0	22.0	29.2	32.7
Sets	10.6	15.0	27.5	46.9
Recites numbers	4.6	11.3	24.9	59.2
Recognizes numbers	6.4	17.1	27.9	48.7
Sorts	4.7	11.7	20.8	62.8
Directional language	6.9	17.3	32.2	43.6
Problem-solving	23.5	24.6	29.0	22.9

Appendix B

Assessment Measures

LITERACY					
Early Reading		**Not Yet**	**Beginning**	**Developing**	**Demonstrating**
CLL5.4a	Prior to reading, uses prior knowledge, story title and pictures to make predictions about story content.		With educator support, can identify what characters may be doing in a story based on the illustrations. When an educator points to a picture on a page and says, "What do you think the characters are going to do now? They just put on their bathing suits and are near the water," the child may say, "They are going swimming."	With less educator support, can make predictions about what will happen next based on illustrations. When an educator turns the page and says, "What do you think will happen now that they have their bathing suits on?" a child may say, "They are going to go swimming."	Without educator support, can make predictions about what will happen based on prior knowledge and illustrations. A child may say, "They are wearing bathing suits. They are going to go swimming."
CLL5.4b	Retells familiar stories.		Can recall key events and people but may miss some details and require some educator support to follow a clear beginning, middle, and end.	Tells stories with a clear beginning, middle, and end with some details missing that may require educator prompts.	Tells stories with detail that cover the beginning, middle, and end without educator prompts.
CLL5.4c	Discusses books or stories read aloud and can identify characters and setting in a story.		With educator support, can identify either characters or setting.	With educator support or with a story they have read multiple times, can identify both the characters and the setting.	Without educator support, can identify characters and a setting.

CLL5.4d	Makes real-world connections between stories and real-life experiences.		Requires educator support. An educator may say, "It looks like the child in this story is sad. Have you ever felt sad?" A child may respond "yes" and recall a time they also felt sad.	Connects simple things from the story to their life. A child may say, "I have blue sneakers, too!"	Without educator support, recalls a familiar experience they had with the experience in a story. A child may say, "I once went to the zoo with my mom like the boy in the story, but I didn't see a tiger—I saw monkeys."
CLL5.4e	Develops an alternate ending for a story.		A child may choose a different ending based on options provided by an educator.	Without support, a child comes up with a unique ending for a story that may or may not be connected to the rest of the story.	Without support, a child comes up with a unique ending for a story that is connected to the rest of the story.
CLL6.4a	Listens and differentiates between sounds that are the same and different.		Without educator support, can identify and differentiate between a few consonant sounds.	With educator support, can identify and differentiate between some vowel and consonant sounds.	Without educator support, identifies and differentiates between some vowel and consonant sounds.
CLL6.4b	Identifies and produces rhyming words.		Can produce rhymes together with educator support during songs.	Can fill in a missing word in a song or sentence that rhymes without educator support.	Generates a rhyming word without educator support or decides that a word does or does not rhyme.
CLL6.4c	Isolates the initial (beginning) sounds in words with adult guidance.		Sings songs that isolate the beginning sounds.	Identifies two words that have the same initial sound.	With educator prompts, can produce the beginning sound in a word.
CLL6.4d	Segments sentences into individual words.		With educator support, can identify one or two words in a sentence.	Without educator support, can identify one or two words in a sentence.	Without educator support, can identify multiple words or all words in a sentence.
CLL6.4e	Segments words into syllables.		Claps along to syllables during songs with educator support.	Claps along or identifies the first syllables of rhyming words without educator support.	Claps along or sounds out a few or all of the syllables of individual words without educator support.

CLL6.4f	Manipulates and blends sounds (phonemes) with adult guidance.		With educator support, can change the initial sound of words to create new words during songs or rhymes.	With some educator support, can change the initial or ending sounds of words to create new words during songs or rhymes.	With some educator support or no educator support, can change the initial or ending sounds of words to create new words during songs or rhymes.
CLL7.4a	With prompting and support, recognizes and names some upper/lowercase letters of the alphabet.		With educator support, recognizes uppercase letters of the alphabet in their name.	With educator support, recognizes and identifies uppercase and lowercase letters of the alphabet in their name.	With some or no educator support, recognizes and identifies 5 to 10 uppercase or lowercase letters of the alphabet.
CLL8.4a	Demonstrates interest in different kinds of literature, such as fiction and nonfiction books and poetry, on a range of topics.		With educator prompts or guidance, flips through different kinds of books.	With peers, explores different types of books.	Independently seeks out different kinds of books.
CLL8.4b	Understands that letters form words. Understands that words are separated by spaces in print.		With educator support, identifies letters within words and can track words separated by spaces in print.	With some educator support, identifies letters within words and can track words separated by spaces in print.	Without educator support, identifies letters within words and can track words separated by spaces in print.
CLL8.4c	With prompting and support, tracks words from left to right, top to bottom and page to page.		With educator support, can point to a few words in a row on a page with one simple sentence.	With educator support, can point to a few words in a row on a page with multiple sentences and follow along.	With little or no educator support, can track a few words in a row on a page.
CLL8.4d	Recognizes and reads environmental print.		With educator support, identifies letters in their name in environmental print.	Without educator support, identifies letters in their name in environmental print.	Without educator support, recognizes familiar or common words in environmental print, such as *stop*.

CLL8.4e	Identifies the front, back, top, and bottom of a book. Points to the title of familiar books or stories and where to begin reading a story.		With educator support, can identify front and back and hold a book with the cover right side up.	Holds a book with the cover right side up and can identify front and back. May need educator support to identify a title and where to begin reading.	Without educator support, can identify a familiar book and flip through from front to back, holding it in the correct position.
CLL9.4a	Draws pictures and copies letters and/or numbers to communicate.		Makes lines, dots, or other markings that may symbolize numbers, letters, or images.	With educator support, makes numbers, letters, or images.	Without educator support, writes numbers, letters, or images.
Early Writing		**Not Yet**	**Beginning**	**Developing**	**Demonstrating**
CLL9.4b	Uses writing tools.		Holds writing tools with their whole hand and makes movements with wrist.	With educator support, uses a three-finger grip and makes movements more fluidly.	Without educator support, uses a three-finger grip.
CLL9.4c	Uses writing for a variety of purposes.		Makes lines, dots, or other markings that may symbolize numbers, letters, or images.	With educator support, makes numbers, letters, or images.	Without educator support, writes numbers, letters, or images.
CLL9.4d	Writes some letters of the alphabet.		Traces, connects the dots, or, with heavy educator support, writes letters.	With educator guidance, freely writes letters of the alphabet.	Without educator support, can write some letters of the alphabet.

MATH					
Number and Quantity		**Not Yet**	**Beginning**	**Developing**	**Demonstrating**
CD-MA1.4a	Recites numbers up to 20 in sequence.		Can recite some numbers in sequence independently but may miss some numbers, say numbers out of sequence, or stop before they get to 20.	With less educator support, can recite most numbers up to 20 in sequence. A child might say, ". . . 11, 12, 14," and after being reminded by the educator that 13 comes next in the sequence, the	Without educator support, can recite numbers 0–20 in order.

				child can recite the remaining numbers in order.	
CD-MA1.4b	Recognizes numerals and uses counting as part of play and as a means for determining quantity.		Can recognize some numerals and uses counting during play but may not count as a means of determining quantity.	Can recognize numerals and counts as part of play. With educator support, can integrate counting as a means for determining quantity into play experiences.	Without educator support, recognizes numerals and counts to determine quantity during play experiences.
CD-MA1.4c	Matches numerals to sets of objects with the same number, 0–10.		With educator support, can count sets of objects and find the matching numeral 0–10.	With less educator support, can match numerals to sets of objects from 0–10.	Without educator support, can match numerals to sets of objects (from 0–10).
CD-MA1.4d	Describes sets as having more, less, same as/equal.		With educator support, can count two sets of objects and describe the relationship between the sets using more, less, or same/as or equal.	With less educator support, can count two sets of objects and describe the relationship between the sets using more, less, or same/as or equal.	Without educator support, can count two sets of objects and describe the relationship between the sets using more, less, or same/as or equal.
CD-MA1.4e	Quickly recognizes and names how many items are in a set of up to four items.		With educator support, can quickly count and identity how many items are in a set of up to four items.	With less educator support, can quickly identify how many items are in a set of up to four items.	Without educator support, can quickly identify how many items are in a set of up to four items.
CD-MA1.4f	Tells numbers that come before and after a given number up to 10.		With educator support, can identify the number that comes before or after a given number within the numbers 0–10.	With less educator support, can identify the number that comes before or after a given number within the numbers 0–10.	Without educator support, can identify the number that comes before or after a given number within the numbers 0–10.

CD-MA2.4a	Matches two equal sets using one-to-one correspondence and understands they are the same.		With educator support, counts two equal sets of objects using one-to-one correspondence.	Counts two equal sets of objects using one-to-one correspondence and, with educator support, understands that the two sets are equal.	Without educator support, counts two equal sets using one-to-one correspondence and recognizes that the two sets are equal.
CD-MA2.4b	Counts at least 10 objects using one-to-one correspondence.		Requires educator support to count objects using one-to-one correspondence.	Can count less than 10 objects using one-to-one correspondence. With less educator support (such as lining objects up or counting with the child), can count 10 objects using one-to-one correspondence.	Without educator support, can count 10 or more objects using one-to-one correspondence.
CD-MA2.4c	Practices combining, separating, and naming quantities.		With educator support, can combine and separate groups of objects and count the quantity of object sets.	With less educator support, can combine and separate groups of objects and identify the quantity of different object sets.	Without educator support, can combine, separate, and identify the quantity of different object sets.
CD-MA2.4d	Describes data from classroom graphs using numerical math language.		With educator support, can explore and talk about data from classroom graphs.	With less educator support, can talk about data from classroom graphs, using numerical math language (number, quantity, more/less, etc.).	Without educator support, can talk about data from classroom graphs, using numerical math language (number, quantity, more/less, etc.).
CD-MA2.4e	With adult guidance and when counting, understands and can respond with the last number counted to represent quantity (cardinality).		Can count a group of objects and, with educator support, recognize that the final number counted represents the quantity of the group (cardinality).	With less educator support, can count a group of objects and understand the final number counted represents the quantity of the group (cardinality).	With minimal educator support, can count a group of objects and understand that the final number counted represents the quantity of the group (cardinality).
Measurement and Comparison		**Not Yet**	**Beginning**	**Developing**	**Demonstrating**

CD-MA3.4a	Uses mathematical terms to describe experiences involving measurement.		With educator support, can use mathematical terms to describe measurement experiences (heavier/lighter, longer/shorter, bigger/smaller, etc.).	With less educator support, can use mathematical terms to describe measurement experiences (heavier/lighter, longer/shorter, bigger/smaller, etc.).	Without educator support, can use mathematical terms to describe measurement experiences (heavier/lighter, longer/shorter, bigger/smaller, etc.).
CD-MA3.4b	Compares objects using two or more attributes, such as length, weight, and size.		With educator support, can compare two objects using one attribute (heavier/lighter, longer/shorter, bigger/smaller, etc.).	With less educator support, can compare two objects using two or more attributes (heavier/lighter, longer/short, bigger/smaller, etc.).	Without educator support, can compare two objects using two or more attributes (heavier/lighter, longer/shorter, bigger/smaller, etc.).
CD-MA3.4c	Uses a variety of techniques and standard and non-standard tools to measure and compare length, volume (capacity) and weight.		With educator support, can measure an object's length, volume (capacity), and weight using nonstandard measurement tools.	With less educator support, can measure an object's length, volume (capacity), and weight using nonstandard measurement tools and standard measurement tools (a ruler, a scale, measuring tape, etc.).	Without educator support, can measure the length, volume (capacity), and weight or two or more objects using nonstandard measurement tools and standard measurement tools (a ruler, a scale, measuring tape, etc.) and then compare those two objects based on those attributes.
CD-MA3.4d	Associates and describes the passage of time with actual events.		With educator support, can describe a simple timeline of actual events ("We ate lunch, then we played outside, then we took a nap.").	Can describe a simple timeline of actual events but may get the order of some events incorrect.	Without educator support, can describe a more detailed timeline of actual events ("First, I woke up at home. Then, I came to school to play. Then, we played outside, and after school, I got ice cream before I ate dinner.").

CD-MA4.4a	Independently orders objects using one characteristic and describes the criteria used.		With educator support, can order a small group of objects based on a characteristic (i.e., size, shape, color).	Without educator support Can order objects based on one characteristic (i.e., size, shape, color).	Without educator support, orders objects based on one characteristic (i.e., size, shape, color) and describes the criteria they used to order items.
CD-MA4.4b	Sorts and classifies objects using one or more attributes or relationships.		With educator support, can sort a small group of objects based on a one attribute or relationship (i.e., size, shape, pattern, type).	With some educator support, can sort objects based on one or more attributes or relationships (i.e., size, shape, pattern, type).	Without educator support, can sort a small group of objects based on one attribute. Independently, or with some educator support, can sort objects based on more than one attribute or relationship (i.e., size, shape, pattern, type).
CD-MA4.4c	Creates and extends simple, repeating patterns.		With educator support, can extend a simple, repeating pattern (ABAB, ABCABC, etc.).	Can extend a simple, repeating pattern. With educator support, can create a simple, repeating pattern (ABAB, ABCABC, etc.).	Without educator support, can create and extend simple, repeating patterns (ABAB, ABCABC, etc.)
Geometry and Spatial Thinking		**Not Yet**	**Beginning**	**Developing**	**Demonstrating**
CD-MA5.4a	Uses appropriate directional language to indicate where things are in their environment - positions, distances, order.		Requires educator support to follow directions using directional language (position, distances, order).	With less educator support, can follow directions using directional language (position, distance, order).	Without educator support, can give and follow directions using directional language (position, distance, order).
CD-MA5.4b	Uses deliberate manipulation and describes process for fitting objects together.		Independently, may use trial and error to make things fit without deliberate actions to manipulate objects. Requires educator support to deliberately manipulate	With less educator support, can practice techniques for deliberate actions to fit things together.	Without educator support, can manipulate objects to fit things together and describe or talk about their process.

			objects to fit things together.		
CD-MA6.4a	Recognizes and names common two-dimensional and three-dimensional shapes, their parts, and attributes.		With educator support, can recognize and name common two- or three-dimensional shapes.	Can name common two- or three-dimensional shapes but needs educator support to identify the parts and attributes of shapes (sides, vertices, etc.).	Without educator support, can identify common two- and three-dimensional shapes and share details about their parts and attributes (sides, vertices, etc.).
CD-MA6.4b	Combines simple shapes to form new shapes.		Manipulates and plays with shapes without recognizing any new shapes that are formed.	With some educator support, can combine two or more simple shapes to make a new shape and names the new shape.	Without educator support, can combine two or more simple shapes to make a new shape and names the new shape.
Mathematical Reasoning		**Not Yet**	**Beginning**	**Developing**	**Demonstrating**
CD-MA7.4a	Estimates using mathematical terms and understands how to check the estimate.		With educator support, can make an estimate about the attributes of an object or a set of objects and use a mathematical strategy to check their estimate.	With less educator support, can make an estimate about the attributes of an object or a set of objects and use a mathematical strategy to check their estimate.	Without educator support, can make an estimate about the attributes of an object or a set of objects and use a mathematical strategy to check their estimate.
CD-MA7.4b	Uses simple strategies to solve mathematical problems and communicates how he/she solved it.		With educator support, can identify real-life mathematical problems and think through how to use math to solve those problems.	With less educator support, can identify real-life mathematical problems and think through how to use math to solve those problems.	Without educator support, can identify real-life mathematical problems, think through how to use math to solve those problems, and tell an adult how they came up with a solution.

CD-MA7.4c	Uses reasoning skills to determine the solution to a mathematical problem and communicates why.		With educator support, can use reasoning skills to solve math problems and communicate how they solved the problem using mathematical terms.	With less educator support, can solve simple math problems, communicate how they solved the problem using mathematical terms, and demonstrate how they came up with the solution.	Without educator support, can solve simple math problems using reasoning skills, communicate how they solved the problem using mathematical terms, and demonstrate how they came up with the solution.