



FORMATIVE CHILD ASSESSMENT (FCA) LITERATURE SUPPORT SUMMARY

I. Definition, Description, and Example of Element

Child assessments help gather information about children's progress in order to understand and support their learning and development (Williams-Appleton et al., 2008). As applied in early childhood education, the term *assessment* generally implies the intention to provide a rich description of the ways in which young children act, think, and learn (Bowman et al., 2001). Assessments in early childhood can have many different purposes and can be conducted using various instruments. The National Education Goal Panel describes four main purposes for early care and education (ECE) assessments: 1) to promote learning and development of individual children; 2) to identify children with special needs and health conditions for intervention purposes; 3) to monitor trends in programs and evaluate program effectiveness; and 4) to obtain benchmark data for accountability purposes at the local, state, and national levels (Shepard et al., 1998). Assessments are considered formative when making decisions about promoting children's learning are based on the information gleaned through observations and interactions (Gipps, 1994; Torrance, 2001 from Dunphy et al., 2010).

In the Implementation Development Map (IDM), the Formative Child Assessment (FCA) Element measures policies, supports, infrastructure, and data collection at the state and local level that support the use of formative child assessments in pre-K programs. Research supports teachers' use of multiple sources of formative assessment data over regular intervals of time to learn more about children's development so teachers are able to provide timely, appropriate instructional support (Ford et al., 2013; McAfee & Leong, 2011). The use of formative child assessment data supports teaching by providing teachers with information so they are able to tailor their instruction to support children's individual learning needs (Riley-Ayers, 2014). The administration of formative child assessments for dual language learners (DLL) and children with special needs requires training and support to teachers and staff so that both the assessments and the interpretation of assessment data are appropriate (culturally and developmentally responsive) and individualized. It is also recommended that education staff share formative child assessment information with families and collaborate with families to support their children's learning.

The FCA Element has two types of indicators: infrastructure and implementation. Equitable infrastructure indicators focus on state systems, policies, and practices that support high-quality pre-K. The infrastructure indicators appear at the beginning of the Element and are labeled as policy (e.g., established in policy and statewide standards), supports (e.g., dedicated resources), and data (e.g., data collection standards and protocols and data use). Equitable implementation indicators focus on the degree to which high-quality pre-K practices occur at the program level and who is benefitting. These indicators require active data collection based upon a representative sample to ensure that all subpopulations are progressing and experiencing the benefits of improvement efforts. Below we list the infrastructure and implementation indicators that make up the FCA Element.

FCA1. FCA Requirements

State requires programs to use formative child assessments (FCAs) that include all three of the following



characteristics:

- Valid and reliable
- Comprehensive across all domains and development
- Aligned to state learning and development standards

FCA2. FCA and special populations (DLLs and special needs)

State requires programs to have a process in place so that FCA data on children who are DLLs and children with developmental delays and disabilities is valid and not misinterpreted due to language or cultural barriers. The state provides clear guidelines on how to do this (e.g., using interpreters or linguistically appropriate assessments) and verifies directly or indirectly that programs comply.

FCA3. Formative Child Assessment Resources

State provides resources (e.g., funding, guidance, etc.) to support teachers in the implementation of formative child assessments.

FCA4. Formative Child Assessment Data

State verifies that programs collect FCA data, and programs use it to:

- Inform improvement plans
- Track progress
- Communicate data to families

In addition, the state uses the data to guide decisions for technical assistance and resource allocation to programs.

FCA5. FCA Data Collection for Equity Goals

The state's efforts to understand and address inequity include ongoing data collection, disaggregation of data, active discussions, data-driven decision-making, action planning, implementing, assessing implementation, and refining as needed. The state specifically collects data to understand and address the following four components:

- FCAs are conducted using reliable and valid methods for all children.
- Data are used to inform instruction, especially instruction that supports the learning and development of children who are Dual Language Learners and children with developmental delays or disabilities.
- All teachers have access to trainings on fidelity of implementation of assessment instruments, as well as trainings on bias when collecting and interpreting data.
- Teachers involve parents in child's education by sharing formative assessment data and seeking guidance from parents when creating individualized instruction plans.

FCA6. Formative Child Assessment Requirements Implementation

Programs use valid and reliable formative child assessment measures that are comprehensive across all domains and aligned to state learning and development standards.



**FCA7. Formative Child Assessment of special population (DLLs and special needs)**

Implementation Programs have a process in place to ensure that formative child assessments of children from special populations are valid and accurate.

FCA8. Formative Child Assessment Classroom Use

Programs use formative child assessment data for classroom quality improvement.

FCA9. Formative Child Assessment Program Use

Programs use formative child assessment data to plan professional development and CQI work.

FCA10. Formative Child Assessment Training

Teachers are trained and, when appropriate, certified in the FCAs they implement.

FCA11. Formative Child Assessment Supports

Programs have supports (e.g., written materials, in-person or online training, teacher collaboration groups) for FCA implementation available to teachers.

FCA12. Family Engagement that is Culturally and Linguistically Appropriate

Teachers regularly share children's progress with families in culturally and linguistically sensitive ways and in their preferred language.

II. Formative Child Assessment Literature Process Overview and Summary

To understand the existing literature support and identify the literature gaps and limitations for each of the IDM indicators, we conducted a systematic literature search and checked with experts for relevant sources to support the various indicators of FCA. More details of the general review process conducted across all elements can be found in the [IDM Evidence Review Document](#). For the FCA element, 15 key phrases were identified and explored. Out of these initial phrases, nine key phrases retrieved relevant results. The list of all sources that yielded relevant results based on the nine key phrases and expert recommendations, along with six key phrases that did not yield relevant results, can be found in Appendix A.

Once the literature search for the FCA Element was completed, we reviewed the quantity and rigor of the literature supporting each indicator and computed what we termed the Literature Support Index (LSI). The LSI is calculated for each indicator based on seven components:

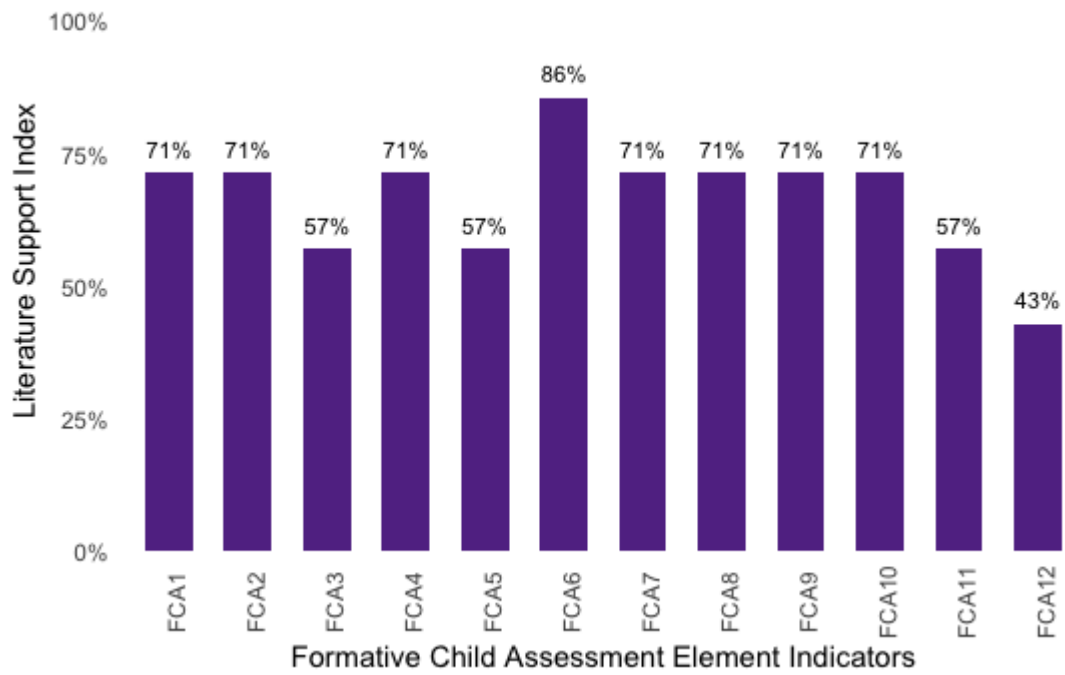
1. at least three peer-reviewed articles;
2. at least one study with no more than two limitations;
3. at least one study at national or state level;
4. at least one study that uses experimental or quasi-experimental design;
5. at least two studies that use representative sampling;
6. support from at least one national research organization; and
7. support from at least one national policy organization.

The LSI is expressed as a percentage of the above seven criteria that are satisfied for a particular indicator. For more information about the rationale for the LSI and how it is calculated can be found in the IDM Evidence Review Document Figure 1 summarizes the LSI for the FCA Element indicators.



Figure 1

Overall Summary of FCA Literature Support Index

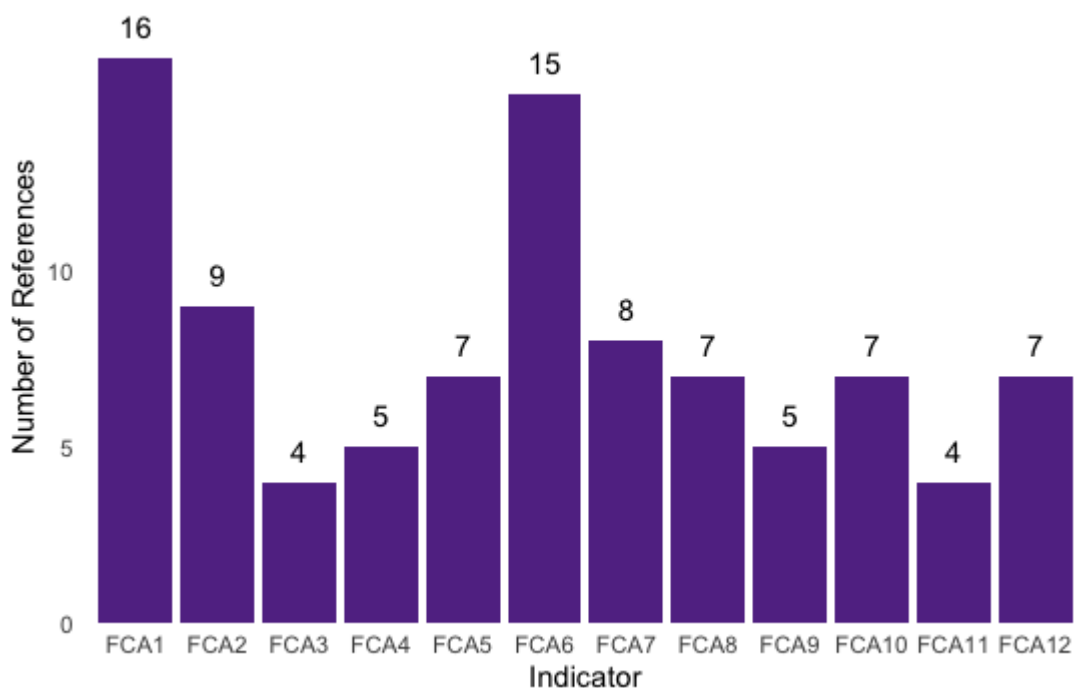




While Figure 1 combines aspects of both rigor of the literature as well as quantity supporting each indicator, Figure 2 presents solely the quantity of evidence for each indicator. Figure 2 shows that FCA indicators 1 and 6 are supported by a larger number of sources than the rest of the indicators. We hope that this type of analysis can help state teams understand where there are gaps in research and potential directions for future studies.

Figure 2

FCA Quantity of Evidence by Indicator

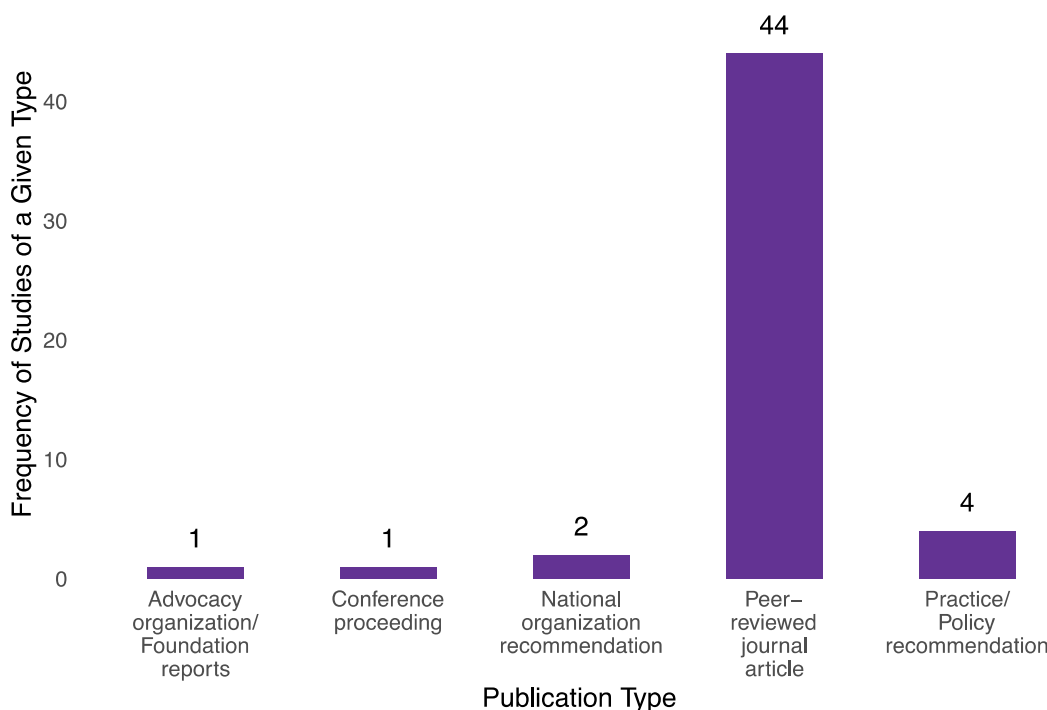




To understand more about the nature of the literature that supports the Element, Figure 3 lists the types of publications used as evidence for the indicators of the FCA Element; the vast majority of the sources are articles from peer-reviewed journals (44).

Figure 3

FCA Evidence by Publication Type



To provide more context on these metrics, we also examined other information related to the relevant article, such as the number of citations, impact factor, publishers behind articles, amount of supporting literature by year published, methodological limitations of studies used as supportive evidence for the FCA indicators, and the scale at which the studies supporting IDM FCA indicators operated. Based on this further analysis, FCA1 is heavily supported by sources that are published in more rigorous journals and used more recent articles, with seven sources published between 2000 and 2005, but twice as much (14) supporting sources published in the last five years. This is in contrast to indicators FCA3, FCA5, and FCA6, where evidence comes from journals that are either too new to have an impact factor or are not considered very relevant in the early learning science community.

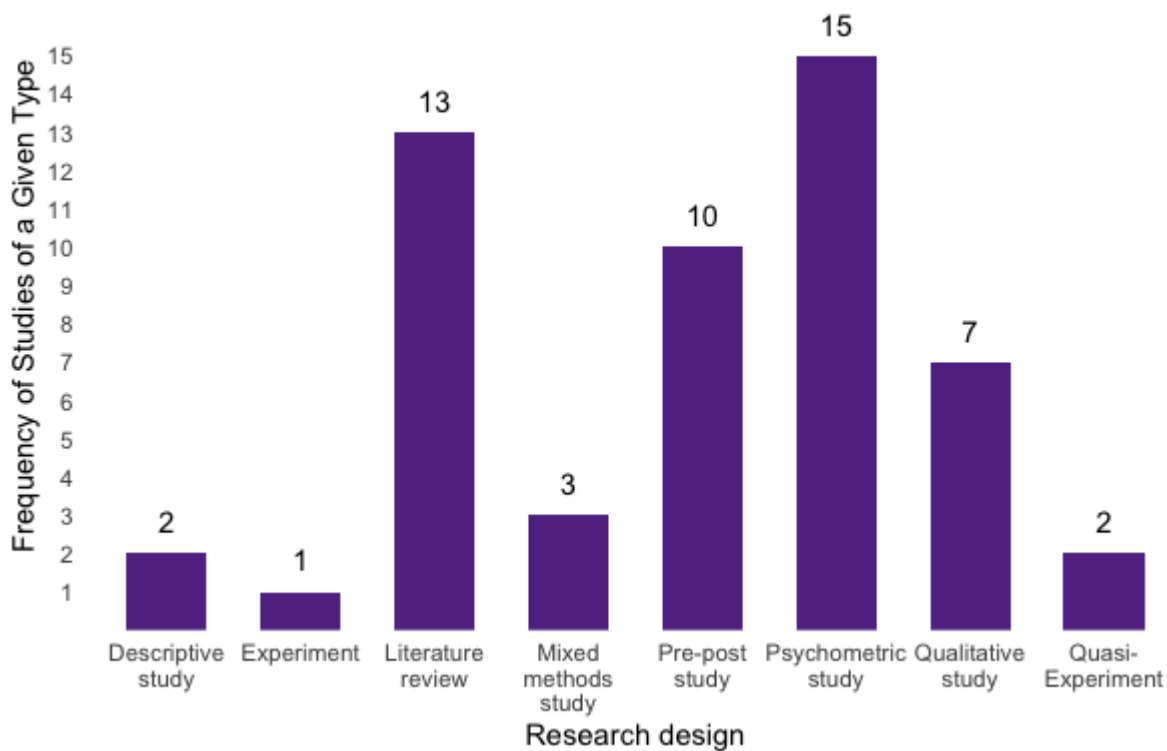




In addition to types of publications, Figure 4 summarizes the research design used in the sources supporting the FCA Element. The most common type of research designs represented in the FCA literature scan (15) involved psychometrics of assessment tools with heavy emphasis on TS GOLD, Child Observation Record, and Desired Results Developmental Profile. Other common study design approaches were pre-post association studies (10), literature reviews (12), and qualitative studies (7). Quasi-experiments and experiments were not highly represented in the types of literature that undergird the FCA Element (2 and 1, respectively).

Figure 4

FCA Summary of Research Design





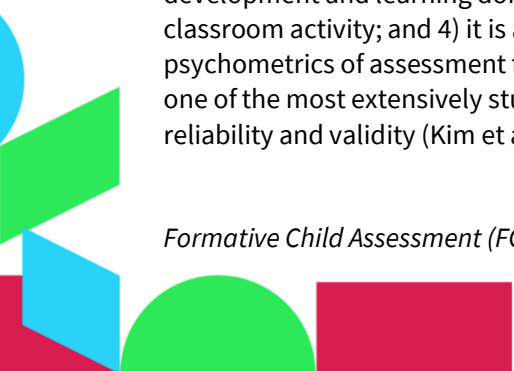
III. Summary of FCA Literature Supporting Indicators: Current Practices and Challenges

This section provides a summary of the literature supporting each indicator, including the current practices and challenges. Because some of the infrastructure and implementation indicators cover the same topics but at the state and program level, we have grouped our summary together to reflect the overlap in the literature for these indicators.

Formative Child Assessment Infrastructure Indicators (state level)	Formative Child Assessment Implementation Indicators (classroom and program level)
FCA1. FCA Requirements State requires programs to use formative child assessments (FCAs) that include all three of the following characteristics: <ul style="list-style-type: none"> • Valid and reliable • Comprehensive across all domains and development • Aligned to state learning and development standards 	FCA6. Formative Child Assessment Requirements Implementation Programs use valid and reliable formative child assessment measures that are comprehensive across all domains and aligned to state learning and development standards.

Position statements by the National Association for the Education of Young Children (NAEYC) describe the importance of assessing children's progress and program effectiveness in coordinated, connected, and continuous ways as it relates to child outcomes (Duran et al., 2013; Fantuzzo et al., 2008), compared to on-demand tests that are disconnected and separate from children's experiences (NAEYC, 2003). Data collected by the Build Initiative & Child Trends (n.d.) in the Quality Compendium on 39 states found that 28 state quality initiatives indicated the use of child assessment to guide individualization and curriculum planning. The most widely used formative child assessment tools are Teaching Strategies GOLD® (Teaching Strategies LLC., n.d.) followed by the High/Scope Preschool Child and Toddler Child Observation Record (The Build Initiative & Child Trends, n.d.).

It is crucial to explore the psychometrics of any assessment tool (Snow & Van Hemel, 2008). As previously mentioned, Teaching Strategies GOLD is the most widely used formative child assessment tool available in Head Start (Isaacs et al., 2015 from Kim, 2016) and state-funded pre-K (Schilder & Carolan, 2014 from Kim, 2016). Teaching Strategies GOLD is considered to be an authentic early childhood assessment tool because of key features, such as: 1) data collection from multiple and varied data sources at multiple times; 2) assessment of all development and learning domains; 3) children are assessed by their own teacher during the course of everyday classroom activity; and 4) it is a curriculum-embedded assessment. As seen in Figure 4 above, studies on psychometrics of assessment tools are widely represented in the FCA Element, and Teaching Strategies GOLD is one of the most extensively studied formative child assessment tools, with several studies examining its reliability and validity (Kim et al., 2014; Kim & Smith, 2010; Lambert et al., 2015; Lambert et al., 2016; Russo et



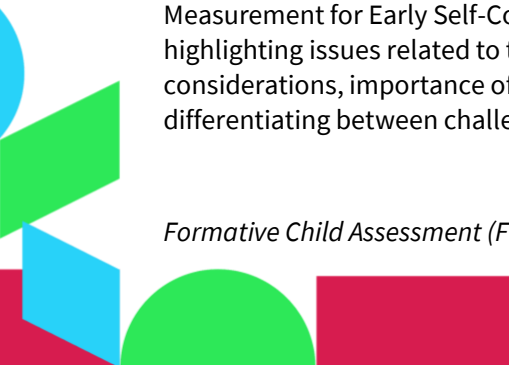


al., 2019). Several studies examined the psychometric properties of other formative child assessment tools, including the Child Observation Record (Barghaus & Fantuzzo, 2014; Fantuzzo et al., 2002; Sekino & Fantuzzo, 2005; Wakabayashi et al., 2019), Desired Results Developmental Profile (DRDP) (Karelitz et al., 2010; Nguyen et al., 2020; Sutter et al., 2017), and the Works Sampling System for Head Start (Meisels et al., 2008). In another study, Fantuzzo et al. (2011) developed an intervention wherein a key component included formative assessments to learn more about children's skills that directly mapped onto state and national Head Start standards. These formative assessments were comprehensive, covering alphabet knowledge, phonemic awareness, vocabulary, print concepts, listening comprehension, mathematics, motor, social-emotional, and approaches-to-learning skills.

Based on the literature above for FCA Indicator 1 and Indicator 6, much of the research focuses on examining reliability, validity, and comprehensiveness of formative child assessments tools such as the Teaching Strategies GOLD, Child Observation Record, and DRDP, with the majority of the studies focusing on Teaching Strategies GOLD. We believe that this is representative of the wide use of Teaching Strategies GOLD for ongoing child assessments in the early learning field.

Formative Child Assessment Infrastructure Indicators (state level)	Formative Child Assessment Implementation Indicators (classroom and program level)
<p>FCA2. FCA and special populations (DLLs and special needs)</p> <p>State requires programs to have a process in place so that FCA data on children who are DLLs and children with developmental delays and disabilities is valid and not misinterpreted due to language or cultural barriers. The state provides clear guidelines on how to do this (e.g., using interpreters or linguistically appropriate assessments) and verifies directly or indirectly that programs comply.</p>	<p>FCA7. Formative Child Assessment of special population (DLLs and special needs) Implementation</p> <p>Programs have a process in place to ensure that formative child assessments of children from special populations are valid and accurate.</p>

In conducting formative child assessments, it is critical to ensure that data on children who are in certain groups such as dual language learners or children with disabilities is valid and not misinterpreted due to language and/or cultural barriers. For instance, recommendations from the National Academies of Science, Engineering, and Medicine (NASEM, 2017) for early childhood staff who assess DLLs include that staff understand the stages of second language acquisition in young children and not compare progress toward meeting program goals with monolingual peers because of differences in the developmental trajectories between these types of language learners. Programs should also establish procedures so they are able to continually assess DLLs' skills in their multiple languages. A few studies have focused on understanding ways to assess whether the measures were culturally and linguistically responsive. For example, McCabe et al. (2000), as part of the Games As Measurement for Early Self-Control (GAMES) project, reviewed measures of children's self-regulation, highlighting issues related to the administration of self-regulation assessments in the field such as pragmatic considerations, importance of standardized administration, need for variability in children's responses, differentiating between challenge and confusion, cultural sensitivity, and recruitment. Another article by Chen





et al. (2009) summarizes evidence-based and recommended practices in assessing young children with disabilities and how authentic assessment approaches using alternative means of communication are more likely to provide helpful information compared to standardized tests, particularly with children who have sensory impairment and multiple disabilities. In a study with a Spanish-speaking preschooler referred for aggression and social deficits, Duran et al. (2013) explored the use of functional behavioral assessment (FBA) and found that providing culturally and linguistically responsive adaptations appeared to help child outcomes, such as reducing problem behaviors during the day. Finally, a psychometric study conducted by Kim et al. (2013) examining measurement equivalence on the Teaching Strategies GOLD found that while the majority of items displayed little or no differential item functioning (DIF), one item pertained to the use of conventional grammar consistently identified as having DIF. The researchers further discuss reviewing this item and the need for high-quality training for teachers who use this tool. They argued that training should incorporate numerous and varied examples of the various ways children with a range of disabilities and language experiences display competencies, so these children are accurately and fairly assessed.

Figure 5 summarizes the types of inequities examined by the sources supporting the FCA Element. The top three types of inequities the various studies focused on were children's race/ethnicity, economic status, and language, with fewer studies focusing on other inequities such as special needs or geographical locations (urban vs rural).



Formative Child Assessment Infrastructure Indicators (state level)	Formative Child Assessment Implementation Indicators (classroom and program level)
<p>FCA4. Formative Child Assessment Data</p> <p>State verifies that programs collect FCA data, and programs use it to:</p> <ul style="list-style-type: none"> ● Inform improvement plans ● Track progress ● Communicate data to families <p>In addition, state uses the data to guide decisions for technical assistance and resource allocation to programs.</p>	<p>FCA8. Formative Child Assessment Classroom Use</p> <p>Programs use formative child assessment data for classroom quality improvement.</p> <p>FCA9. Formative Child Assessment Program Use</p> <p>Programs use formative child assessment data to plan professional development and CQI work.</p>

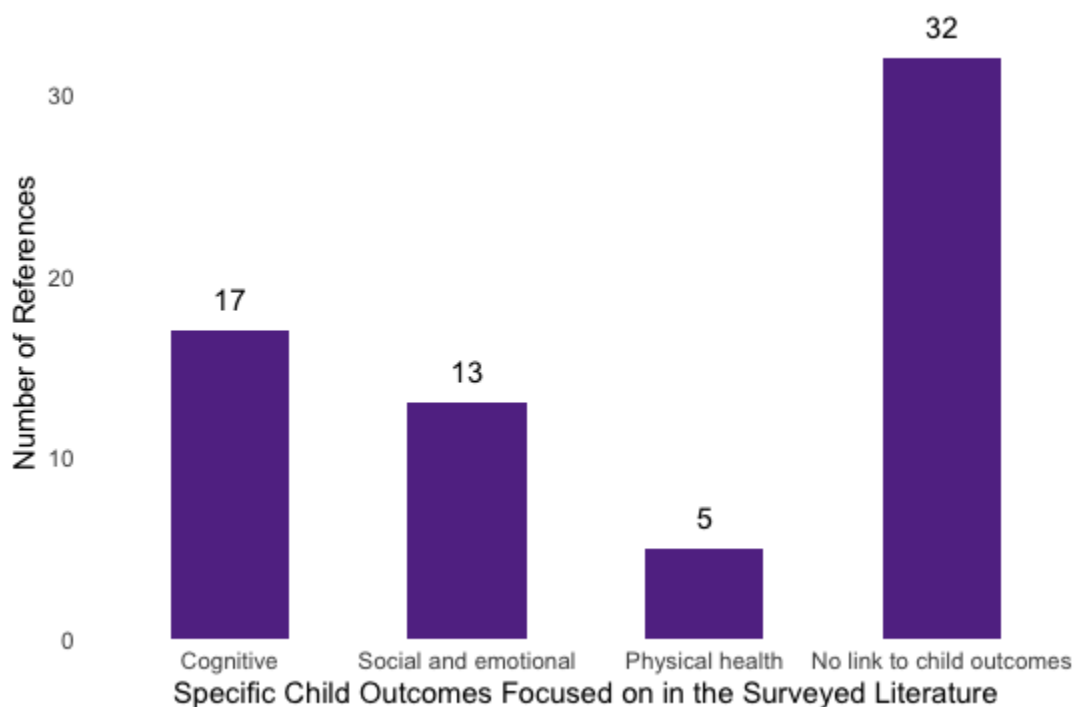
One of the key reasons for using FCA in early childhood is for teachers to use the data to inform instruction at the classroom level. As Figure 5 shows, only a little more than half of the studies examine the outcomes of children who participate in interventions that focus on the development and implementation of formative child assessments (35). The rest of the studies are often the precursors to such studies because the scientific community first has to come to an agreement as to how such an improvement would be measured (32 studies in the bar on the right in Figure 5). The studies that did examine the relationship between the use of formative child assessment and child outcomes looked at the cognitive (17), emotional (13), and health benefits (5). A study by Little et al. (2019) on data use in North Carolina’s pre-K program found that while there was much data collected through developmental screening tools and formative assessment systems, the use of these data for instruction is variable.





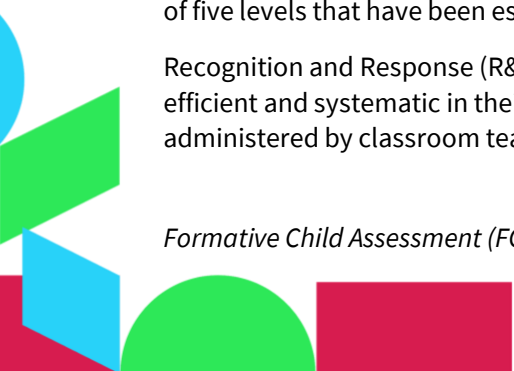
Figure 5

FCA Child Outcomes Studies Examined



There are two studies that use formative child assessments as a key feature in their intervention: Recognition & Response (R&R) study conducted by Buysse et al. (2016) and Evidenced-Based Program for an Integrated Curriculum (EPIC) conducted by Fantuzzo et al. (2011). Fantuzzo et al. (2011) developed integrated check-ins (ICIs) as part of the EPIC curriculum. ICIs are brief assessments of children's skill level that can help teachers monitor children's progress and create a classroom profile of individual student ability to inform instruction. Teachers completed ICIs as part of their routine implementation of the curriculum and repeated ICIs three times throughout the year. These skills directly map onto state standards for early childhood education and the national Head Start indicators, and include alphabet knowledge, phonemic awareness, vocabulary, print concepts, listening comprehension, mathematics, motor, social-emotional, and approaches-to-learning skills. On their EPIC curriculum with 70 Head Start classrooms, children (1,415) were randomly assigned to one of two curriculum programs: EPIC or the Developmental Learning Materials Early Childhood Express, with curricula implemented as standalone programs. Multilevel growth modeling through four direct assessments revealed positive child outcomes, with significant main effects and growth rates in mathematics and listening comprehension favoring EPIC, controlling for demographics and special needs and language status. Both programs produced significant growth rates in literacy. Each skill is assessed across a developmental sequence of five levels that have been established and validated by empirical research (Fantuzzo et al., 2008).

Recognition and Response (R&R) is a multi-component instructional system designed to help teachers be more efficient and systematic in their use of formative assessments (recognition) of key language and literacy skills, administered by classroom teachers on computer tablets to all children periodically throughout the year, and



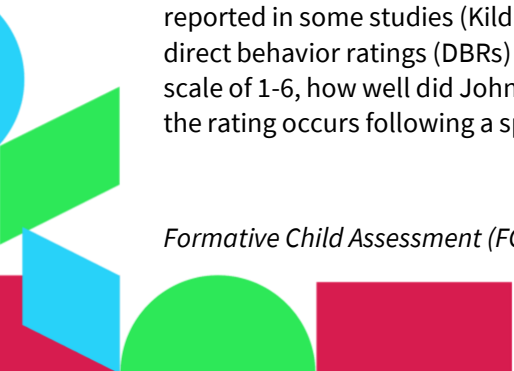


curricular resources in order to differentiate instruction (response) to meet the varying learning needs of children within a typical pre-K classroom (Buyse et al., 2016). Two studies conducted on R&R indicate positive child outcomes for target children who received the small-group lessons, including larger gains on some language and literacy skills than a comparison group consisting of classmates who had lower than average scores at baseline and did not receive the small-group lessons (Buyse et al., 2016).

Overall, studies show the value of formative child assessment data to support instruction and individualized learning goals for children at the classroom level, but the use of these data may be variable. At the state level, pre-K leaders can use the data to understand equity issues in school readiness among all children across the state and to guide a number of decisions, including policies around formative child assessment standards, coaching, and TA for teachers to conduct and make use of such assessments, and allocate resources to programs to support improvement of children's learning and school readiness.

Formative Child Assessment Infrastructure Indicators (state level)	Formative Child Assessment Implementation Indicators (classroom and program level)
<p>FCA3. Formative Child Assessment Resources</p> <p>State provides resources (e.g., funding, guidance, etc.) to support teachers in the implementation of formative child assessments.</p>	<p>FCA10. Formative Child Assessment Training</p> <p>Teachers are trained and, when appropriate, certified in the FCAs they implement.</p> <p>FCA11. Formative Child Assessment Supports</p> <p>Programs have supports (e.g., written materials, in-person or online training, teacher collaboration groups) for FCA implementation available to teachers.</p>

These indicators focus on ensuring that the necessary supports are available for teachers to implement formative child assessments with their students with fidelity and to use assessment data to improve instruction. Training on the use of formative child assessments can facilitate teachers' awareness of the influence their perceptions (Bennett et al., 1993; Burchinal et al., 2011) and classroom contexts (Gallagher & Lambert, 2006; Meisels et al., 2010) have on child appraisals. However, in a literature review conducted by Akers et al. (2015), they found evidence to suggest that although teachers might recognize the value of ongoing assessment, they do not consistently collect or use ongoing assessment data to tailor their instruction. Research studies have found that teacher-based observational assessment is more subjective than standardized measures (Cabell et al., 2009) and there is the possibility for greater variability (Kilday et al., 2012), particularly for students identified as dual language learners and children with developmental delays (Joseph et al., 2020). Some researchers have questioned whether teacher reports represent actual child differences or other factors such as teacher variability or classroom context (Lambert et al., 2014; Waterman et al., 2012). In their study on Teaching Strategies GOLD, Lambert et al. (2014) found error variance ranged from 16% to 25%, considerably lower than reported in some studies (Kilday et al., 2012). Another study by Chafouleas et al. (2007) discusses the use of direct behavior ratings (DBRs) that refer to a rating process similar to that of a behavior rating scale (e.g., "On a scale of 1-6, how well did Johnny pay attention?"), yet it is similar to systematic direct observation given that the rating occurs following a specified shorter period of time (Chafouleas et al., 2007). Results suggested that a





fairly substantial proportion of measurement variance was attributable to the different raters, and that the four raters varied in their mean level of ratings within and across ratings. Given that work suggests teachers' professional experiences are significant predictors of agreement level (Joseph et al., 2020), additional research is needed on providing comprehensive professional development related to the collection and use of ongoing assessment data, and the use of technology-driven support, which seems to be more effective than no professional development (Akers et al., 2015).

Given that studies suggest the importance of ongoing formative child assessment, and the importance of training and professional development for teachers to complete these types of assessments, states have a role and responsibility in ensuring equitable and sustainable professional development opportunities are available for teachers to develop the skills, knowledge, and understanding of how to conduct assessments and how to use the resulting assessment data.

Formative Child Assessment Infrastructure Indicators (state level)

FCA5. FCA Data Collection for Equity Goals

The state's efforts to understand and address inequity include ongoing data collection, disaggregation of data, active discussions, data-driven decision-making, action planning, implementing, assessing implementation, and refining as needed. The state specifically collects data to understand and address the following four components:

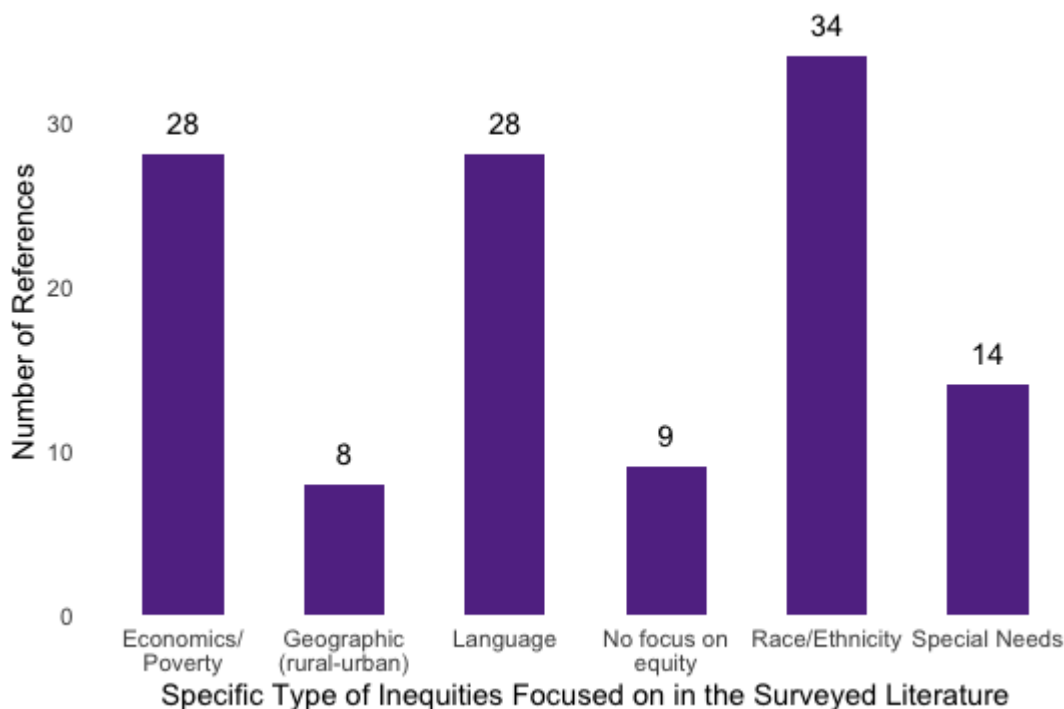
- FCAs are conducted using reliable and valid methods for all children.
- Data are used to inform instruction, especially instruction that supports the learning and development of children who are Dual Language Learners and children with developmental delays or disabilities.
- All teachers have access to trainings on fidelity of implementation of assessment instruments, as well as trainings on bias when collecting and interpreting data.
- Teachers involve parents in child's education by sharing formative assessment data and seeking guidance from parents when creating individualized instruction plans.





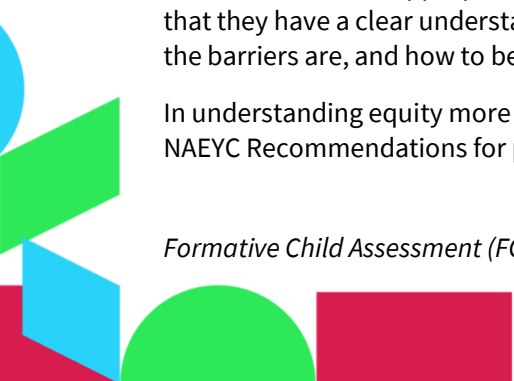
Figure 6

FCA Inequities of Focus in the Literature



In line with the framework of targeted universalism (Powell et al., 2019) used to guide the development of the IDM, equity indicators in each Element highlight the importance of ongoing data collection, the disaggregation of data, and the use of data for decision-making, action planning, and assessing implementation. This supports the five steps of targeted universalism (Powell et al., 2019), where once a universal goal is established (Step 1), and there is information about the performance of the general population relative to the universal goal (Step 2), the performance of different groups can be identified (Step 3), further analysis can be done to understand the structures that support or impede each group for achieving the universal goal (Step 4), and targeted strategies for each group can be developed and implemented to reach the universal goal (Step 5). As seen in Figure 6, based on the literature, the most common types of inequities studied were related to race/ethnicity (33), with many studies measuring the implementation of child assessment and its relationship with child outcomes across various racial and ethnic groups. The next most common type of inequities examined related to socio-economic background, with studies again sharing information about family demographics and making comparisons across various family income groups. Inequities related to language were also fairly common (27), and also the focus of FCA Indicator 2 and FCA Indicator 7. We encourage state leaders to collect formative child assessment data in appropriate, valid, and reliable methods for all children represented in their population so that they have a clear understanding of how children from different backgrounds and regions are faring, what the barriers are, and how to best mitigate barriers.

In understanding equity more specifically within the context of formative child assessments, we examine the NAEYC Recommendations for policymakers for advancing equity in early childhood education (2020). The





recommendations encourage the use of authentic assessments that are developmentally, culturally, and linguistically appropriate, in addition to being valid and reliable and used for a purpose consistent with the goals of the tool. While formative child assessments focus on the regularity and timing of conducting assessments, authentic assessments are concerned with the way tools capture data—relying on information that can be obtained in a child’s natural environments (Neisworth & Bagnato, 2004). Further, an important distinction about authentic assessments is that they are not artificial or decontextualized, with the sole purpose of making the child demonstrate a behavior, characteristics often found in conventional psychometric standardized tools. Instead, educators who use authentic assessments look for an example/evidence of the skill of interest during a child’s daily activities at home and in settings outside the home, such as school, the supermarket, or church. Authentic assessments thus focus on real behaviors that have a functional importance rather than standardized item content that has little instructional value (Neisworth & Bagnato, 2004).

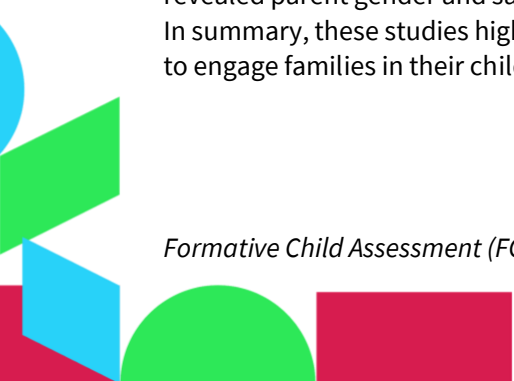
Ensuring formative child assessments are conducted in authentic ways will increase the likelihood that the data collected about children are developmentally, culturally, and linguistically appropriate (as further discussed in FCA 2 and FCA 7). Additionally, authentic assessments provide opportunities for families to be more engaged in their children’s education by having teachers work with families to gather meaningful data about their children rather than relying on data from the teacher in the school setting.

Formative Child Assessment Implementation Indicators (classroom and program level)

FCA12. Family Engagement that is Culturally and Linguistically Appropriate

Teachers regularly share children's progress with families in culturally and linguistically sensitive ways and in their preferred language.

Families are important partners in their children’s development. Indicator 12 for the FCA Element focuses on examining how teachers regularly share children’s progress with families in culturally and linguistically sensitive ways and in their preferred language. Although research is limited in this area, several qualitative studies describe families as important partners in the collection and interpretation of child data (Akers et al., 2015; Chen et al., 2009; Michael-Luna et al., 2013) as well as illustrate teachers’ and families’ challenges in having these discussions, especially if they do not come from the same cultural and linguistic backgrounds (Cheatham et al., 2013; Mendez, 2010). The Family Involvement Questionnaire (FIQ; Fantuzzo, 2013, 2004, 2000) includes questions related to the communication between families and teachers about child progress and experiences. One of the three factors is home–school conferencing, which focuses on asking questions related to how early childhood educators better understand how families perceive communication and potential communication methods to share child progress and information. For example, McWayne et al. (2008) used the FIQ to examine the relationships between family demographics and level of satisfaction with school contact, and regression analysis highlighted differences in mothers’ and fathers’ school-based involvement. Further multilevel analyses revealed parent gender and satisfaction as the most salient predictors of involvement at the level of the family. In summary, these studies highlight the need to establish shared expectations and to explore meaningful ways to engage families in their children's development to maximize the educational outcomes of all children.





IV. Future Directions and Limitations

ECE teachers' use of ongoing assessment has not been extensively researched, and the research that does exist lacks scope and rigor (Akers et al., 2015). As mentioned, there are a limited number of articles that focus on training teachers, and no articles that specifically studied the role of programs and the state to provide resources (e.g., funding, guidance, written materials, in-person or online training, teacher collaboration groups, etc.) to support teachers in the implementation of formative child assessments.

In addition, there is a gap in studies that focus on types of training for teachers on formative assessment tools as well as considering other sources of data. For example, Reeves et al. (2017) studied how integrating mobile devices into a pre-K curriculum using informal feedback from students affects students' academic achievement. Results of the ANCOVA revealed significantly higher phonological awareness and mathematics measures for the iPad class, suggesting that integrating mobile learning in content-specific areas using informal student feedback effectively increases early childhood education students' academic achievement. Best practices for integrating mobile learning to enhance student engagement are discussed.

In reviewing the limitations of the studies used to support the FCA Element, we found that most studies (apart from the studies on Teaching Strategies GOLD) that support the FCA Element reach their conclusions from small samples and scale. Still, 11 studies demonstrated their conclusions using a nationally representative sample (and an additional five on a statewide sample) of children, which for that reason would be the strongest evidence—indicating that not only does the conclusion hold in small elite pilots but also in large scale implementation (that will be most relevant to the users of IDM). Nevertheless, these study limitations need to be taken into account, with the anticipation that more state pre-K studies can provide insight into how state policies and practices can best adapt these findings to support implementation in the classrooms.





References

- Akers, L., Grosso, P. D., Atkins-Burnett, S., Monahan, S., Boller, K., Carta, J., & Wasik, B. A. (2015). Tailored teaching: The need for stronger evidence about early childhood teachers' use of ongoing assessment to individualize instruction. In *Mathematica Policy Research Reports* (Mathematica Policy Research Reports). Mathematica Policy Research. <https://ideas.repec.org/p/mpr/mpres/1cd21071996f4a7cba4260e8b18fea8a.html>
- Bagnato, S. J., Goins, D. D., Pretti-Frontczak, K., & Neisworth, J. T. (2014). Authentic assessment as "best practice" for early childhood intervention: National consumer social validity research. *Topics in Early Childhood Special Education*, 34(2), 116–127. <https://doi.org/10.1177/0271121414523652>
- Bagnato, S. J., McLean, M., Macy, M., & Neisworth, J. T. (2011). Identifying instructional targets for early childhood via authentic assessment: Alignment of professional standards and practice-based evidence. *Journal of Early Intervention*, 33(4), 243–253. <https://doi.org/10.1177/1053815111427565>
- Barghaus, K. M., & Fantuzzo, J. W. (2014). Validation of the preschool child observation record: Does it pass the test for use in Head Start? *Early Education and Development*, 25(8), 1118–1141.
- Beneke, M. R., Newton, J. R., Vinh, M., Boone Blanchard, S., & Kemp, P. (2019). Practicing inclusion, doing justice: Disability, identity, and belonging in early childhood. *Zero to Three*, 39(3), 26–34.
- Brown, J., & Rolfe, S. A. (2005). Use of child development assessment in early childhood education: Early childhood practitioner and student attitudes toward formal and informal testing. *Early Child Development and Care*, 175(3), 193–202. <https://doi.org/10.1080/0300443042000266240>
- Buyse, V., Peisner-Feinberg, E., & Burchinal, M. (2012). Recognition & response: Developing and evaluating a model of RTI for pre-k. In *Society for Research on Educational Effectiveness*. Society for Research on Educational Effectiveness. <http://eric.ed.gov/?id=ED530413>
- Chafouleas, S. M., Christ, T. J., Riley-Tillman, T. C., Briesch, A. M., & Chanese, J. A. M. (2007). Generalizability and dependability of direct behavior ratings to assess social behavior of preschoolers. *School Psychology Review*, 36(1), 63–79.
- Cheatham, G. A., & Ostrosky, M. M. (2013). Goal setting during early childhood parent-teacher conferences: A comparison of three groups of parents. *Journal of Research in Childhood Education*, 27(2), 166–189. <https://doi.org/10.1080/02568543.2013.767291>
- Chen, D., Rowland, C., Stillman, R., & Mar, H. (2009). Authentic practices for assessing communication skills of young children with sensory impairments and multiple disabilities. *Early Childhood Services: An Interdisciplinary Journal of Effectiveness*, 3(4), 323–338.
- Dunphy, E. (2010). Assessing early learning through formative assessment: Key issues and considerations. *Irish Educational Studies*, 29(1), 41–56. <https://doi.org/10.1080/03323310903522685>
- Durán, L. K., Bloom, S. E., & Samaha, A. L. (2013). Adaptations to a functional behavior assessment with a Spanish-speaking preschooler: A data-based case study. *Education & Treatment of Children*, 36(1), 73–95. <https://doi.org/10.1353/etc.2013.0009>
- Edyburn, K. L., Dowdy, E., DiStefano, C., Bertone, A., & Greer, F. (2020). Measurement invariance of the English and Spanish BASC-3 behavioral and emotional screening system parent preschool forms. *Early Childhood Research Quarterly*, 51, 307–316. <https://doi.org/10.1016/j.ecresq.2019.12.002>
- Espinosa, L. M. (2020). Addressing equity in the ECE classroom: equal access and high quality for dual language learners. *Getting It Right*, 21.
- Fantuzzo, J. W., Gadsden, V. L., & McDermott, P. A. (2011). An integrated curriculum to improve mathematics, language, and literacy for Head Start children. *American Educational Research Journal*, 48(3), 763–793. <https://doi.org/10.3102/0002831210385446>
- Fantuzzo, J., Gadsden, V., Li, F., Sproul, F., McDermott, P., Hightower, D., & Minney, A. (2013). Multiple dimensions of family engagement in early childhood education: Evidence for a short form of the Family Involvement Questionnaire. *Early Childhood Research Quarterly*, 28(4), 734–742. <https://doi.org/10.1016/j.ecresq.2013.07.001>



- Fantuzzo, J., Hightower, D., Grim, S., & Montes, G. (2002). Generalization of the child observation record: A validity study for diverse samples of urban, low-income preschool children. *Early Childhood Research Quarterly*, 17(1), 106–125. [https://doi.org/10.1016/S0885-2006\(02\)00131-X](https://doi.org/10.1016/S0885-2006(02)00131-X)
- Fantuzzo, J., McWayne, C., Perry, M. A., & Childs, S. (2004). Multiple dimensions of family involvement and their relations to behavioral and learning competencies for urban, low-income children. *School Psychology Review*, 33(4), 467–480.
- Fantuzzo, J., Tighe, E., & Childs, S. (2000). Family Involvement Questionnaire: A multivariate assessment of family participation in early childhood education. *Journal of Educational Psychology*, 92(2), 367–376. <https://doi.org/10.1037/0022-0663.92.2.367>
- Federal Statute, Head Start Act of 2007, Child Assessment, Head Start Performance Standards. §1302.12-1302 Subpart C, Subchapter B. (2007).
- Joseph, G., Soderberg, J. S., Stull, S., Cummings, K., McCutchen, D., & Han, R. J. (2020). Inter-rater reliability of Washington State's kindergarten entry assessment. *Early Education and Development*, 31(5), 764–777. <https://doi.org/10.1080/10409289.2019.1674589>
- Karelitz, T. M., Parrish, D., Yamada, H., & Wilson, M. (2010). Articulating assessments across childhood: The cross-age validity of the Desired Results Developmental Profile-Revised. *Educational Assessment*, 15(1), 1–26. <https://doi.org/10.1080/10627191003673208>
- Kim, K. (2016). Teaching to the data collection? (Un)intended consequences of online child assessment system, “Teaching Strategies GOLD.” *Global Studies of Childhood*, 6(1), 98–112. <https://doi.org/10.1177/2043610615627925>
- Kim, D.-H., Lambert, R. G., & Burts, D. C. (2013). Evidence of the validity of “Teaching Strategies GOLD®” assessment tool for English language learners and children with disabilities. *Early Education and Development*, 24(4), 574–595.
- Kim, D.-H., Lambert, R. G., & Burts, D. C. (2014). Validating a developmental scale for young children using the Rasch model: Applicability of the Teaching Strategies GOLD® assessment system. *Journal of Applied Measurement*, 15(4), 405–421.
- Kim, D.-H., & Smith, J. (2010). Evaluation of two observational assessment systems for children's development and learning. *NHSA Dialog*, 13(4), 253–267.
- Kriener-Althen, K., Newton, E. K., Draney, K., & Mangione, P. L. (2020). Measuring readiness for kindergarten using the Desired Results Developmental Profile. *Early Education and Development*, 0(0), 1–25. <https://doi.org/10.1080/10409289.2020.1743160>
- Lambert, R. G., Kim, D.-H., & Burts, D. C. (2014). Using teacher ratings to track the growth and development of young children using the Teaching Strategies GOLD® assessment system. *Journal of Psychoeducational Assessment*, 32(1), 27–39.
- Lambert, R. G., Kim, D.-H., & Burts, D. C. (2015). The measurement properties of the Teaching Strategies GOLD® assessment system. *Early Childhood Research Quarterly*, 33, 49–63. <https://doi.org/10.1016/j.ecresq.2015.05.004>
- Lambert, R. G., Kim, D.-H., & Burts, D. C. (2016). The association between the Teaching Strategies GOLD® assessment system and direct assessments of child progress. *AERA Online Paper Repository*.
- Little, M., Cohen-Vogel, L., Sadler, J., & Merrill, B. (2019). Data-driven decision making in early education: Evidence from North Carolina's pre-k program. *Education Policy Analysis Archives*, 27(18). <https://files.eric.ed.gov/fulltext/EJ1210325.pdf>
- Ma, X., Shen, J., Krenn, H. Y., Hu, S., & Yuan, J. (2016). A meta-analysis of the relationship between learning outcomes and parental involvement: During early childhood education and early elementary education. *Educational Psychology Review*, 28(4), 771–801. <https://doi.org/10.1007/s10648-015-9351-1>
- Mccabe, L. A., Hernandez, M., Lara, S. L., & Brooks-Gunn, J. (2000). Assessing preschoolers' self-regulation in homes and classrooms: Lessons from the field. *Behavioral Disorders*, 26(1), 53–69. <https://doi.org/10.1177/019874290002600106>
- McWayne, C., Campos, R., & Owsianik, M. (2008). A multidimensional, multilevel examination of mother and father involvement among culturally diverse Head Start families. *Journal of School Psychology*, 46(5), 551–573. <https://doi.org/10.1016/j.jsp.2008.06.001>
- Meisels, S. J., Xue, Y., & Shablott, M. (2008). Assessing language, literacy, and mathematics skills with Work Sampling for Head Start. *Early Education and Development*, 19(6), 963–981. <https://doi.org/10.1080/10409280801971890>



- Michael-Luna, S. (2013). What linguistically diverse parents know and how it can help early childhood educators: A case study of a dual language preschool community. *Early Childhood Education Journal*, 41(6), 447–455. <https://doi.org/10.1007/s10643-013-0574-9>
- Neisworth, J. T., & Bagnato, S. J. (2004). The mismeasure of young children: The authentic assessment alternative. *Infants & Young Children*, 17(3), 198–212. <https://doi.org/10.1097/00001163-200407000-00002>
- Nguyen, T., Reich, S. M., Jenkins, J. M., & Abedi, J. (2020). Psychometric validation and reorganization of the desired results developmental profile. *Journal of Psychoeducational Assessment*, 38(3), 369–388. <https://doi.org/10.1177/0734282919859804>
- Paris, D., & Alim, H. S. (2014). What are we seeking to sustain through culturally sustaining pedagogy? A loving critique forward. *Harvard Educational Review*, 84(1), 85–100. <https://doi.org/10.17763/haer.84.1.982l873k2ht16m77>
- Powell, J., Menendian, S., & Ake, W. (2019). *Targeted universalism: Policy & practice*. Haas Institute for a Fair and Inclusive Society. <https://belonging.berkeley.edu/targeteduniversalism>
- Recommendations for Public Policymakers | NAEYC. (n.d.). Retrieved November 9, 2020, from <https://www.naeyc.org/resources/position-statements/equity/recommendations-policy>
- Russo, J. M., Williford, A. P., Markowitz, A. J., Vitiello, V. E., & Bassok, D. (2019). Examining the validity of a widely-used school readiness assessment: Implications for teachers and early childhood programs. *Early Childhood Research Quarterly*, 48, 14–25. <https://doi.org/10.1016/j.ecresq.2019.02.003>
- Sekino, Y., & Fantuzzo, J. (2005). Validity of the child observation record: An investigation of the relationship between core dimensions and social-emotional and cognitive outcomes for Head Start children. *Journal of Psychoeducational Assessment*, 23(3), 242–260.
- Snow, C. E., & Van Hemel, S. B. (2008). Early childhood assessment. Why, what, and how? National Research Council of the National Academies Report.
- Stein, A., Freel, K., Hanson, A. T., Pacchiano, D., & Eiland-Williford, B. (2013). The Educare Chicago research-program partnership and follow-up study: Using data on program graduates to enhance quality improvement efforts. *Early Education and Development*, 24(1), 19–41. <https://doi.org/10.1080/10409289.2013.739542>
- Sutter, C., Ontai, L. L., Nishina, A., Conger, K. J., Shilts, M. K., & Townsend, M. S. (2017). Utilizing the desired results developmental profile as a measure of school readiness: Evaluating factor structure and predictors of school readiness. *Early Child Development and Care*, 187(9), 1433–1445. <https://doi.org/10.1080/03004430.2016.1174933>
- Wakabayashi, T., Claxton, J., & Smith, E. V. Jr. (2019). Validation of a revised observation-based assessment tool for children birth through kindergarten: The COR advantage. *Journal of Psychoeducational Assessment*, 37(1), 69–90. <https://doi.org/10.1177/0734282917732491>
- Waterman, C., McDermott, P. A., Fantuzzo, J. W., & Gadsden, V. L. (2012). The matter of assessor variance in early childhood education—Or whose score is it anyway? *Early Childhood Research Quarterly*, 27(1), 46–54. <https://doi.org/10.1016/j.ecresq.2011.06.003>





Appendix A

Table 1 Formative Child Assessment Literature Scan Summary of Relevant Articles

Key word or phrase	# Articles for initial abstract review based on inclusion criteria	# Articles for 2nd abstract review with exclusion criteria	# Articles passed full article review	Article citation
Child observation record	31	5	5	Barghaus and Fantuzzo, 2014 (FCA1, 6); Fantuzzo et al., 2002 (FCA1, 6); Meisels et al., 2008 (FCA1, 6); Sekino and Fantuzzo, 2005 (FCA1, 6); Wakabayashi et al., 2019 (FCA1)
Culturally responsive	90	14	1	McCabe et al., 2000 (FCA2, 7)
Data-driven decision-making	55	18	1	Little et al., 2019 (FCA3, 8, 9,)
Desired results developmental profile	10	4	4	Karelitz et al., 2010 (FCA1, 6); Kriener-Althen et al., 2020 (FCA2); Nguyen et al., 2020 (FCA1, 6); Sutter et al., 2017 (FCA1, 6)
Home-school conferencing	6	3	2	Fantuzzo et al., 2000 (FCA12); McWayne et al., 2008 (FCA12)
Family engagement	343	19	5	Ma et al., 2016 (FCA12); Cheatham and Ostrosky, 2013 (FCA12); Fantuzzo et al., 2013 (FCA12); Fantuzzo et al., 2004 (FCA12); Michael-Luna, 2013 (FCA12)
Formative assessment	145	13	5	Buyse et al., 2012 (FCA8); Chafouleas et al., 2007 (FCA3, 6, 10, & 11); Chen et al., 2009 (FCA2, 7); Fantuzzo et al., 2011 (FCA4, 8, & 9); Stein et al., 2013 (FCA9)
Linguistically responsive	26	8	2	Duran et al., 2013 (FCA2, 7); Edyburn et al., 2020 (FCA10)
TS GOLD	38	8	8	Kim and Smith, 2010 (FCA 1, 6); Kim et al., 2013 (FCA2, 7); Russo et al., 2019 (FCA1);

Key word or phrase	# Articles for initial abstract review based on inclusion criteria	# Articles for 2nd abstract review with exclusion criteria	# Articles passed full article review	Article citation
				Kim, 2016 (FCA4); Lambert et al., 2016 (FCA1, 6); Lambert et al., 2015 (FCA1, 6); Lambert et al., 2014 (FCA2, 3, 10, & 11) Kim et al., 2014 (FCA1, 6)
Expert recommendation			15	Akers et al., 2015 (FCA3, 10, & 11); Bagnato et al., 2011 (FCA8); Bagnato et al., 2014 (FCA5); Beneke et al., 2019 (FCA5); Brown and Rolfe, 2005 (FCA10); Dunphy, 2010 (FCA1); Espinosa, 2020 (FCA5); Joseph et al., 2020 (FCA5); NAEYC Equity Statement, 2020 (FCA5); Neisworth et al., 2004 (FCA5); Paris and Alim, 2014 (FCA5); Powell et al., 2019 (FCA5); Snow and Van Hemmel, 2008 (FCA1, 6, 8, 9, & 10); Waterman et al., 2012 (FCA3, 10, & 11); Head Start Performance Standards, 2007 (FCA6, 7)
Data analysis	30	0	0	N/A
Information utilization	30	0	0	N/A
Evidence-based	32	0	0	N/A
High/Scope	59	5	0	N/A: Articles were relevant for HQT and RBC

Key word or phrase	# Articles for initial abstract review based on inclusion criteria	# Articles for 2nd abstract review with exclusion criteria	# Articles passed full article review	Article citation
Work sampling system	10	2	0	N/A: Relevant article taken out because sample was K-3 not pre-K (did not meet our criteria)
Individualization/ Individualized support	66	4	0	N/A
Total	971	103	47	

Acknowledgement

This work was created as part of the Partnership for Pre-K Improvement, a collaboration between the [Alliance for Early Success](#), [Cultivate Learning](#), and [Start Early](#), with support from the [Bill & Melinda Gates Foundation](#).

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

© 2021 University of Washington. All rights reserved.