SRI International

ENHANCE Final Report

Grant Award Number R324A090171

SRI Project Number: P19032

Prepared for:

U.S. Department of Education
Institute of Education Sciences
555 New Jersey Ave., NW
Washington, DC 20208-5500

Attention: Leontyne B. Minor
Director, Grants Administration

Prepared by:

Lauren Barton
Donna Spiker
Kathleen Hebbeler
Cornelia Taylor
Cyndi Williamson

SRI International

Contents

Appendices

A   COS Materials

B   Selected Resources Related to Outcomes Constructs

C   State Requirement Materials from the U.S. Department of Education

D   ENHANCE Provider Survey

E   Provider Survey Data for EI and ECSE

F   Psychometric Characteristics of the BDI-2 and Vineland-II

G   Crosswalks of BDI-2 and Vineland-II to Child Outcomes

H   ENHANCE Child and Family Information Form

I   Additional Descriptive Information about Longitudinal Sample

J   Child Assessments Study: Regressions Predicting COS Ratings

K   Levels of 7-Point COS Ratings with Mean Scores on BDI-2, Vineland-II, and ABILITIES Index at Program Entry in the Child Assessments Study

L   Team Decision-Making Study Project Developed Forms

M   Team Decision-Making Study: Regressions Predicting COS Ratings

N   ENHANCE Project Dissemination

Executive Summary

The Office of Special Education Programs (OSEP) in the U.S. Department of Education requires that states report annually on the outcomes of children under 5 years of age who received services under the auspices of the Individuals with Disabilities Education Act (IDEA). This report presents the findings from a series of studies conducted to examine the validity of the child outcomes data produced through the Child Outcomes Summary (COS) process.

The COS process was designed to provide a structure and a rubric for local teams (e.g., teachers, parents, early interventionists, therapists) to synthesize multiple sources of information about a child's functioning across settings and situations. In the COS process, teams apply criteria to identify how the child's functioning compares with age-expected functioning using a 7-point metric for each of three outcomes: (1) Children have positive social-emotional skills (including social relationships), (2) Children acquire and use knowledge and skills (including early language/communication [and early literacy[1]]), and (3) Children use appropriate behaviors to meet their needs. The COS ratings are used to determine children's progress between entry into and exit from early intervention (EI) or Part C programs (programs for children birth to 3 years old) or early childhood special education (ECSE) or Part B preschool programs (programs for 3- through 5-year-olds) (Hebbeler & Kahn, 2014). States aggregate this information and report percentages of children in each of five progress categories with regard to each of the three outcomes.

The COS method of outcomes data collection is used by 75% of states and territories (42 of 56) for reporting on EI and 63% of states and territories (37 of 56) for reporting on ECSE. Given the widespread use of the COS to generate data for state and national accountability, a significant need exists to examine the validity and reliability of the data produced through this process.

This project, funded by the U.S. Department of Education's Institute of Education Sciences, addressed three objectives related to the validity of COS data: (1) conduct a program of research to examine the validity of ratings generated by the COS process and identify conditions that lessen validity, (2) revise the COS form and supporting materials based on study findings, and (3) identify a series of validity analyses that can feasibly be conducted in states to allow each state to examine the validity of its own COS data on an ongoing basis. Given that an objective of EI and ECSE programs is to enhance child outcomes and that the research focus was on understanding how to enhance the quality of COS outcomes data, this project was called ENHANCE.

Drawing on the guidance of national experts in measurement and statistics, the research team conducted four studies to examine 16 claims about the validity of the COS process. The four studies examined the content of the COS, the rating process, the constructs addressed, the

---

[1]   Applies to ECSE only.

relationship of COS ratings to other variables, and the consequences associated with COS use. The studies involved surveying providers who were implementing the COS process, comparing COS ratings with the results of two child assessments, examining COS team meetings through the use of videotapes, and analyzing state data from state EI and ECSE programs using the COS. Studies 1 through 3 were carried out in 35 local programs, 19 EI and 16 ECSE programs in eight states.

- In **Study 1 (provider survey),** the research team used online surveys to examine providers' perceptions of the COS, the process, the training and support they received, and its impact on practice. Through the survey, providers also shared information about their knowledge of the outcomes, experiences implementing the COS process, and indicated the degree of difficulty they had in applying the rating criteria during team decisions ($n = 856$ providers, 472 EI providers, 302 ECSE providers, and 82 in both EI and ECSE).

- In **Study 2 (comparison with child assessments study**), the research team examined the relationship between COS ratings and two standardized child assessment tools, the Battelle Developmental Inventory, Second Edition (BDI-2) and the Vineland Adaptive Behavior Scales, Second Edition (Vineland-II). Data were collected at two time points with children in local EI and ECSE programs from seven states. The team examined relationships among the COS, BDI-2, and Vineland-II approaches as well as the differences among them with regard to producing data for OSEP ($n = 153$ children with entry data and $n = 70$ with exit data). This study involved comparing the COS ratings with the BDI-2 and Vineland-II assessments to learn the extent to which children were classified in the same progress category by each of these approaches.

- In **Study 3 (team decision-making study**), the researchers examined fidelity of COS implementation, including accuracy of the ratings, through analysis of videos of teams making decisions about COS ratings ($n = 113$, 63 EI teams and 50 ECSE teams). The researchers examined both structural and process elements associated with team rating decisions and completing the COS form. The videos were coded to determine whether or not the team reached a rating that was consistent with the available evidence presented and examined team members' knowledge and application of knowledge of the content of the three outcomes, the rating criteria for each of the 7 points on the rating scale. Videos also were coded for accurate use of child development behaviors and age expectations during team discussions and other indicators of a quality COS process.

- In **Study 4 (extant state data study**), the research team examined multiple validity claims using extant data by analyzing datasets from 18 state programs (9 EI, 9 ECSE) using the COS. Extant data from state programs, and national data compiled by the Early Childhood Outcomes center and the Early Childhood Technical Assistance Center using the COS were analyzed to examine specific claims shown Exhibit 1.4.

Across the four studies, evidence was drawn for the validity of claims concerning the Child Outcomes Summary process.

- The findings were used to test 13 validity claims.
- Support was found for 11 of the claims, and the evidence was mixed for the other two.
- Two additional claims were supported with evidence collected outside the study.
- One claim could not be tested because no data were collected.

Key study findings related to the validity of the COS addressed providers' understanding of the content of the outcomes, the patterns in the ratings, the relationship of the ratings to other assessment tools, expected patterns for subgroups, and the lack of negative consequences of implementing the COS process. Based on the evidence across the series of validity claims, we concluded that the COS process was capable of producing valid data for accountability and program improvement purposes.

On the basis of the findings from these studies, we expected that the resulting data would indicate any necessary revisions to the COS and the supporting materials such as the instructions and the training materials (see Section 8 of the final report for a description of these implications). During the ENHANCE project, the research team members used what they were learning from ENHANCE and from their ongoing interactions with states to improve the materials available to states to support the COS process. Suggested revisions of resources are presented and many of these resources are now under development. Implications for future research also are presented including a need to know more about the conditions under which the validity of the data is enhanced and those which lessen it.

Section 1 — Background

The Office of Special Education Programs (OSEP) in the U.S. Department of Education requires that states report annually on the outcomes of children under 5 years of age who received services under the auspices of the Individuals with Disabilities Education Act (IDEA). This report presents the findings of a series of studies that were conducted to examine the validity of the child outcomes data produced through the Child Outcomes Summary (COS) process. The COS process involves the child's team reviewing multiple sources of information and applying criteria to derive a rating of the child's functioning. The COS method of outcomes data collection is used by 75% of the states and territories (42 of 56) for reporting on early intervention (EI or Part C programs, programs for children birth to 3 years old) and 63% of the states and territories (37 of 59) for reporting on early childhood special education (ECSE or Part B preschool programs, programs for 3- through 5-year-olds) (Hebbeler & Kahn, 2014). Given the widespread use of the COS to generate data for state and national accountability, a significant need exists to examine the validity and reliability of the data produced through this process.

The Need for Good Data on Outcomes for Young Children with Disabilities

About 20 years ago, policymakers and funders began to recognize the importance of holding programs and agencies accountable for intended results (Hogan, 2001; Morley, Vinson, & Hatry, 2000; Osbourne & Gaebler, 1992). Previously, programs had reported data about the service process, such as number of hours of services or the number of clients. This new focus created a need for data on outcomes, which involved documenting children's changes in functioning resulting from services. Accountability for results extends to programs serving children with disabilities, including those below school age (President's Commission on Excellence in Special Education, 2002). Since the passage of the Government Performance Results Act of 1993 (GPRA) (Senate Committee on Governmental Affairs, 1993), each federal agency has been required to develop a strategic plan and report on outcomes data annually (Senate Committee on Governmental Affairs, 1993). This requirement applies to OSEP, which is responsible for implementing IDEA and monitoring the state agencies that administer Part C and Part B Preschool programs. However, for many years, the only data the states and OSEP could provide about these programs were process data (e.g., number of children and families served).

A critical event in the development of accountability systems for Part C and Part B Preschool occurred in 2002 when the federal Office of Management and Budget (OMB) implemented a new review process, the Program Assessment Rating Tool (PART). PART findings were to be used for policymaking priorities and budgetary decisions with the expectation that ineffective or poorly performing programs would be reduced or eliminated. Part C and Part B Preschool were included in the first set of federal programs reviewed in 2002. Both programs received a rating of

"Results Not Demonstrated" because they did not have data on outcomes (U.S. Office of Management and Budget, 2006). The PART process considerably increased pressure on OSEP for data on child outcomes by making it explicit that future program funding would be tied to the availability of these data. An initiative of the Bush administration, the PART was discontinued under President Obama, but the policy emphasis on producing and measuring outcomes has continued.

Challenges in Measuring Child Outcomes

Together, GPRA and PART created a pressing need for data on child outcomes, but obtaining these data required building a national measurement system. The collection of data on outcomes that could be aggregated across classrooms, programs, and ultimately states would require common measurement procedures. Barriers to obtaining data on outcomes for young children with disabilities included the lack of agreement about appropriate outcomes; the inadequacies of assessment tools for young children, especially children with disabilities; and the extreme variability in functioning of children served in early intervention (EI) and early childhood special education (ECSE) (Hebbeler & Barton, 2007). In a comprehensive review of the science of early childhood, a joint National Research Council and Institute of Medicine (2000a) committee wrote, "...for more than three decades, researchers and service providers have struggled with both the identification of significant child outcomes and their valid and reliable measurement." A report from the National Research Council (NRC) (2001) elaborated on these issues: "Assessment of young children poses greater challenges than people generally realize. ... There is widespread dissatisfaction with traditional norm-referenced standardized tests, which are based on early 20th century psychological theory" (p. 12). More recently, another NRC report on early childhood assessment reiterated this conclusion (National Research Council, 2008).

The early childhood community has voiced widespread concern about the appropriateness of testing young children for accountability purposes (National Early Childhood Accountability Task Force, 2007). Their concern was based on the high variability in young children's demonstration of their skills in different contexts (e.g., home, child care), when interacting with different individuals (e.g., caregivers versus strangers), and in response to tests that require varied levels of mainstream cultural knowledge and language skill (Meisels, 2006; National Research Council, 2001, 2008). The approach to accountability adopted by the federal government for Head Start proved so controversial that many early childhood experts lobbied to have it removed and an NRC Committee was formed to examine appropriate approaches to accountability for programs for young children (National Research Council, 2008).

The challenges in building meaningful accountability systems for general early childhood programs were even more significant for programs serving young children with disabilities. Few assessment instruments are appropriate for, or have been validated on, the full range of young children with disabilities, and very few have been designed to measure outcomes with this

population for accountability purposes (Hebbeler, Barton, & Mallik, 2008). Although young children with disabilities are assessed to determine eligibility and to measure individual progress in their programs, many different assessment tools are used and progress is examined based on an individualized set of outcomes outlined in the specific plan developed for that child. Therefore, data cannot be aggregated. To systematically collect data on the same set of outcomes for young children with disabilities meant that every state would need to build an entirely new data collection system.

Developing Accountability Systems for Programs for Young Children with Disabilities

Identifying Child Outcomes

OSEP funded the Early Childhood Outcomes (ECO) Center in October 2003 to provide the federal government and states with guidance and leadership on how to build an accountability measurement system for EI and ECSE programs. To determine a set of appropriate child and family outcomes, the ECO Center engaged in a yearlong process that included analysis of existing outcome frameworks, a literature review, and synthesis of themes from stakeholder discussions. Focus groups were held with state EI and ECSE coordinators, researchers, providers, parents of children with disabilities, and other stakeholders.

Several recurring themes emerged. Strong support was voiced for *a single set of outcomes* for both the Part C and Part B Preschool programs. Stakeholders encouraged the use of outcomes that were *sufficiently global* to be relevant to children with all types and severities of disabilities. Global outcomes also would enable states to map the OSEP outcomes to the varied early learning standards that exist in many states. Stakeholders strongly supported making the child outcomes *functional* (i.e., outcomes that have meaning in the child's everyday life). A clear message was that the outcomes should *not be built around traditional child developmental domains* (e.g., cognition, motor skills) because recommended practice is to think about children from a transdisciplinary perspective that focuses more on how children integrate skills across these domains in functional ways (McWilliam, 2004). The stakeholders felt strongly that a domains-based approach to outcomes would reinforce an outdated intervention model of individual therapists addressing skills limited to those covered in their professional disciplines. Stakeholders recognized that deviating from the traditional domain structure would pose challenges for assessment, but they believed it was essential that the outcomes reflect recommended practice rather than be driven by the structure of existing assessment tools.

After a lengthy input-gathering process, the ECO Center recommended a set of outcomes to OSEP for national data collection (Bailey et al., 2006; Early Childhood Outcomes Center, 2005). OSEP accepted the recommendation and adopted three child outcomes, which were similar for both Part C and Part B Preschool. The outcomes areas are:

- Children have positive social-emotional skills (including social relationships).
- Children acquire and use knowledge and skills (including early language/communication [and early literacy[2]]).
- Children use appropriate behaviors to meet their needs.

Exhibit 1.1 provides more detail about these outcomes.

Exhibit 1.1 Additional Information About The Three Child Outcomes

Outcome	Refers to how a child
Children have positive social-emotional skills (including social relationships)	- Initiates and maintains positive social interactions. - Builds and maintains relationships with children and adults. - Interacts in ways that allows them to participate socially in a variety of settings and situations, for example, on the playground, at dinner, at the grocery store, in child care, preschool class, etc. - Understands and follow socially established rules and norms. - Resolves conflicts in socially acceptable and age appropriate ways.
Children acquire and use knowledge and skills (including early language/communication [and early literacy]	- Thinks, reasons, remembers and retrieves information, across a variety of everyday routines and activities - Acquires pre-academic knowledge and readiness skills, such areas as, communication, language, early literacy, pre-mathematics and mathematics - Shows imagination and creativity, and the ability to think symbolically in play. - Combines and uses knowledge to solve problems - Uses vocabulary either through spoken means, sign language, or through augmentative communication devices to communicate in an increasingly complex form.
Children use appropriate behaviors to meet their needs	- Initiates actions to meet physical needs (such as those for comfort, safety, and well-being), as well as psychological needs (such as the need to master and engage with one's environment). - Uses gestures, sounds, words, signs or other means to make wants and needs known to others. - Meets self-care needs (feeding, dressing, toileting, etc.) and/or seeks help when necessary to assist with basic care or other needs. - Follows rules related to health and safety.

---

[2] "and early literacy" applies to ECSE only.

What States Must Report

OSEP requires that each state report annually on the progress of children who recently exited either early intervention or preschool special education. States report on the percentages of children who made different kinds of progress in five categories:

(a) Did not make any progress

(b) Made progress but not sufficient to move closer to same-age peers

(c) Made progress and moved closer to same-age peers

(d) Achieved functioning comparable to same-age peers

(e) Maintained functioning comparable to same-age peers.

Percentages of children in each of these progress categories are reported with regard to each of the three outcomes, resulting in 15 different numbers being reported. Progress is to be measured between program entry and exit for all children who were in the program at least 6 months. OSEP has not and will not mandate that states use a particular assessment instrument to measure outcomes. States also report on summary statements, which are calculated from the percentage of children in the (a) to (e) progress categories above. The summary statements are

1. Of those children who entered the program below age expectations in each Outcome, the percentage who substantially increased their rate of growth by the time they turned 3/6 years of age or exited the program $[(c + d)/(a + b + c + d)]$

2. The percentage of children who were functioning within age expectations in each Outcome by the time they turned 3/6 years of age or exited the program $[(d + e)/(a + b + c + d + e)]$.

In their February 2008 *Annual Performance Reports*, states reported these data to OSEP for the first time for children who exited the program between July 1, 2006, and June 30, 2007. States were first informed of the requirement in August 2005, and in December 2005 they submitted plans about how they would collect data, including data for children entering in the 2005–06 fiscal year. Faced with a very tight timeline, states had to decide on measurement procedures and implement them shortly after OSEP announced the reporting requirement.

The Child Outcomes Summary (COS) Process

The Need for a Summary Tool

The three functional child outcomes have met with widespread acceptance among administrators, providers, and parents around the country.[3] However, the characteristics of the three outcomes that stakeholders felt positively about also are associated with several measurement challenges: Assessment tools are organized around domains rather than the three outcome areas, some tools do not assess functional skills, and no tool has been designed to measure the three outcomes directly (although various tools tap some of the content of one or

---

[3] See Colorado's website as an example: http://www.cde.state.co.us/resultsmatter/rm_system.htm

more outcome). Given that multiple tools were already being used within the states and no new funding was provided for new assessments, most states wanted to build their accountability systems around existing assessment procedures and allow programs to continue to use different assessments. States were looking for a process that would allow them to aggregate data across programs that were using different assessment tools. To meet these needs, the ECO Center undertook development work exploring alternative ways to capture multiple pieces of information about a child, including the results from different assessments, and produce a single rating for each outcome. Several alternative approaches were developed and shared with various groups of state agency staff and local providers. Their feedback was incorporated into revisions, and the current version of the Child Outcomes Summary process, the focus of this research, was released to states just in time to meet mandated timelines.

Overview of the COS

The COS[4] was designed to provide a structure and a rubric for local teams (e.g., teachers, parents, early interventionists, therapists) to synthesize multiple sources of information on child functioning across settings and situations. The teams apply criteria to identify the appropriate point on a 7-point metric for each of the three outcomes. (Appendix A presents a copy of the COS and instructions for its use, including a decision tree with criteria for rating categories.) Criteria define each point on the scale.

Ratings of 6 or 7 indicate that the child is showing age-appropriate functioning related to the outcome across settings and situations; the 1 to 5 ratings indicate varying degrees of distance from age-appropriate functioning or levels of foundational skills necessary to develop age-appropriate functioning. For example, a 4 on the scale means the child exhibits some age-appropriate functioning but rarely.

In addition to being on an ordinal scale, the ratings are relativistic in that they reflect a comparison of a child's functioning with normative expectations, as needed to address OSEP's reporting requirements. Because the COS ratings are relativistic (nomothetic), rather than idiographic (Singer & Willett, 2003), a rating of 4 at time 1 and again at time 2, for example, indicates the child will have made normative progress but not closed the gap toward age-expected functioning.

The COS process was developed to be used by programs serving young children with disabilities. Thus, rating points were intentionally designed to be sensitive to change among children who are functioning below age expectations (i.e., change that can be qualitatively described with regard to age expectations) and to have limited sensitivity among children whose functioning is considered age appropriate. Nearly all children in the general population would be

---

4   When initially released, the COS process was referred to by many as the COSF (Child Outcomes Summary Form). Use of COS or COS process is preferred to emphasize the significance of the decision-making process rather than the form itself in identifying appropriate ratings of children's functioning.

expected to receive a rating of 6 or 7, whereas children in EI and ECSE programs would be expected to reflect the full 1 to 7 range.

Use of the COS

COS ratings are expected to result from a team decision-making process involving the individuals familiar with the child. For each of the three outcomes, the team is to consider multiple sources of information in reaching a rating. Sources could include formal tools (e.g., norm- and/or criterion-referenced assessment tools) as well as less structured measures (e.g., clinical judgments, case notes, parent observations) representing experiences observing the child's functioning in different situations and with different people. On the basis of its review of multiple sources of information, the team applies the criteria for the rating scale to reach a consensus rating of a child's level of functioning on each of the outcomes across settings and situations. To apply the criteria, the team must consider the extent to which the child's functioning reflects age-appropriate levels or identify the distance of that functioning from age-expected levels based on where in the sequence of development the skills and behaviors observed fall. At exit ratings, the teams also identify whether or not the child made any progress, (that is, learned one or more new skills) since his or her entry into the program.

For OSEP reporting, COS ratings must be completed for each child near program entry and again near program exit. States are required to report only on the progress of children who stayed in the program at least 6 months. Some states have opted to use the COS more frequently to provide better information for program improvement and to reduce missing data at exit for children who exit unexpectedly. A comparison of COS ratings at program entry and exit provides sufficient information to classify a child into one of the five progress categories used for OSEP reporting.

Research Supporting the Assumptions Underlying the COS Process

As the COS process was developed, it reflect several assumptions related to the development of children with disabilities, the overall goal of providing intervention for young children with disabilities, the nature of current assessment tools, the value of multiple sources of information, the utility of judgment-based assessment processes, and the ability of teams to reach consensus on children's functioning. More information about the basis and content of these assumptions is below.

An underlying assumption of the COS process, supported by both research and experiences of practitioners, is that different children display different levels of functioning on the three outcomes (i.e., having positive social skills, acquiring and using knowledge and skills, and using appropriate behavior to meet their needs). It also assumes that these differences can be observed and described. Level of functioning is related to age; older children generally display more sophisticated and complex behaviors than younger children in each outcome area. Furthermore, development and learning in many areas occur in predictable ways, with later skills building on early skills in similar ways for most children.

Active participation in a variety of settings and situations is an overarching goal for children with disabilities and thus is the ultimate goal for providing intervention. Because of this goal, the COS rating was designed to reflect children's proficiencies across a variety of settings and with different people. Current assessment tools provide information organized around domains, and some may not capture functioning across multiple settings. The recognition of the dangers of reaching conclusions about a young child based on a single assessment tool has resulted in many organizations calling for the use of multiple sources of information when assessing young children (National Association for the Education of Young Children & National Association Early Childhood Specialists in State Departments of Education, 2003; National Association of School Psychologists 2005; Neisworth & Bagnato, 2005). Those who know both the child (including functioning in different situations) and the content of the assessment tool are the only people able to determine how well the results from a given assessment accurately reflect that child's current level of functioning on the three outcomes across settings and situations. The team is instructed to consider results for formal assessments in reaching a COS rating, but exact agreement with the results of formal assessment is not expected because of the limitations of currently available assessment tools.

A key component of IDEA service provision is that teams of professionals and the child's parents are to discuss and collectively agree on the child's needs and then develop an individualized plan of services. Research supports the COS assumption that teams also can reach reliable and valid judgments about a child's level of functioning. Highly relevant to these aims is the research on judgment-based assessment (LeLaurin, 1990). In a review of the research foundations for using clinical judgment, Bagnato and colleagues (2006) concluded that clinical judgment produces reliable, valid, and useful information under certain conditions. Important conditions identified included an operational definition of the child's characteristics, a structured format for quantifying those characteristics, use of information from multiple individuals and multiple settings, training in appropriate methods, and a decision-making process based on consensus of multiple individuals. (The COS process includes all of these.)

A rigorous and well-designed set of relevant studies investigated the agreement between interdisciplinary teams rating the functional abilities of children with disabilities using the ABILITIES Index. The ABILITIES Index is a descriptive measure of a child's abilities across nine domains—audition, behavior, intellectual function, limbs, intentional communication, tonicity, physical health, eyes, and structural status (Bailey, Buysse, Simeonsson, Smith, & Keyes, 1995). The results for ratings of 129 children indicated that the 72 team members' level of agreement was high for ratings of abilities and limitations, with nearly 90% of the ratings within 1 point. In a related study, Bailey, Simeonsson, Buysse, and Smith (1993) examined the agreement between ratings made by 130 parents and 125 teachers and specialists. Agreement across the three groups was 86.2%. At a second time point, an average of 34 days later, 90.6% of the ratings were within 1 point of the original rating. Overall, the research demonstrated that raters who vary in expertise and relationship to the child generally agree about the child's abilities.

A recurring question in early childhood has been the extent to which parents and professionals agree in their assessments of the child (Beckman, 1984; Bricker & Squires, 1989; Shaw, Hammer, & Leland, 1991). One question has been whether the appropriate perspective should be one of rater interchangeability (in which case, parents are likely to rate their child differently from professionals and this inconsistency is seen as problematic) versus one in which parents add unique information (as in multicontent, multimethod, multi-informant measurement of latent constructs)(Campbell & Fiske, 1959) to team decisions when pooling ratings. A rigorous study of parent-professional congruence reported adequate reliabilities for judgmental-based assessment outcomes using the System to Plan Early Childhood Services (SPECS) (Suen, Lu, Neisworth, & Bagnato, 1993). SPECS is a comprehensive team-assessment and decision-making battery that converges information from parents and professionals. Suen and colleagues' (1993) sample consisted of 467 kindergarten students (262 were developmentally delayed; 205 were typically developing peers). The reliability of SPECS outcomes in seven of eight scoring and interpretation scenarios was adequate (.71 to .95), with higher reliabilities for scenarios including parents (.92 to .95) and reasonably small standard errors of measurement. Suen and colleagues (1995) concluded that concerns over parent-professional congruence should be abandoned because parents add unique information and because educational policies mandate parent involvement. Bagnato, Matesa, Smith-Jones, and Fevola (2004) concluded that judgment-based assessments that rely on the combined knowledge of parents and professionals are valuable sources of information about child functioning. With regard to COS procedures, all states require parent input be included in the discussion of the child's functioning in keeping with COS procedures, but states vary as to whether or not they routinely include parents as part of the team that decides the rating.

Collectively, the studies of team decision-making suggest that adequate levels of rater reliability are possible. In both the ABILITIES and SPECS studies, if team members were provided with specific guidance and rubric types of procedures for quantifying their judgments, as opposed to no procedures, technical adequacy improved. Also of interest conceptually and empirically is the notion that including parents as team members with the professionals provides information to the process that increases validity and reliability, rather than introducing a source of disagreement relative to the professional members (Suen et al., 1995). The existing research suggests that a judgment-based assessment system such as COS with defined scoring criteria that incorporates multiple sources of information and involves team consensus decision-making holds promise as a tool for state and national accountability systems. One counterclaim that would lessen the validity of COS ratings is that providers will intentionally or unintentionally inflate ratings to make their program look good (Meisels, 2007).

Establishing Validity of the COSF

Approach to Measuring Validity

The Council of Chief State School Officers (CCSSO) publication *A Framework for Examining Validity in State Accountability Systems* provides a helpful approach for validating educational accountability systems (Council of Chief State School Officers, 2004; Fast & Hebbler, 2004). Because these systems entail high-stakes consequences, states and the federal government have an obligation to examine the validity of the systems (CCSSO, 2004). Exactly what validity concepts apply to accountability systems and how those concepts differ from those commonly used with measurement instruments (e.g., tests) are matters of some debate, but there is consensus that examinations of the validity of an accountability system must be anchored to the purposes for which the data are collected (National Research Council, 2008).

The approach to examining the validity of the ratings from the COS process derives from recent work that conceptualizes validation as the process of developing a scientifically sound validity argument (Kane, 1992, 2006; Mislevy, 2006; Moss, Girard, & Haniford, 2006). Neither validity nor reliability are a characteristic of an assessment; rather, they are characteristics of a set of scores derived from an assessment (or in this case, an assessment summary process) (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999; Thompson, 2003). According to the AERA/APA/NCME *Standards*, "Validity refers to the degree to which *evidence and theory* support the interpretation of test scores entailed by proposed uses of test.... The process of validation involves accumulating evidence to provide a sound scientific basis for the use of the tests" (p. 9). The *Standards* go on to note that validity can be addressed by developing a set of propositions that support the proposed interpretation for a set of scores and collecting evidence to examine those propositions. Validation proceeds by developing empirical evidence, examining relevant literature, and conducting logical analyses to examine each of the propositions.

According to the guidelines in the *Standards*, the data from the COS would be considered valid to the extent that the accumulated evidence supports the intended interpretation of the data for the proposed purposes. Validity goals differ at the national, state, and program levels; the purposes for collecting COS data at each of these levels are shown in Exhibit 1.2.

Exhibit 1.2   Validity Goals for COS Data at National, State, and Program Levels

National Level	State Level	Program Level
• To provide data that will enable the U.S. Department of Education and the U.S. Congress to reach sound conclusions about the effectiveness with which states are providing early intervention and early childhood special education services so that appropriate funding decisions can be made.	• To provide data to enable states to reach sound conclusions about the status and progress of children being served in these programs so these data can be used to support funding, program planning, and technical assistance decisions; • To identify local programs that are especially effective, as well as those that need additional assistance in producing good outcomes; and • To identify subgroups of children whose outcomes are especially good, as well as those whose outcomes are not as good as desired, so that decisions can be made about the programmatic changes needed to improve outcomes for all children.	• To produce data that enable programs to examine their overall effectiveness as well as their effectiveness with each segment of the population served so they can make decision about programmatic changes needed to improve outcomes for all children.

If the COS process can produce data that are valid for state decision-making, then the data will be valid when aggregated for federal decision-making. The primary purpose of the data from the COS process is for state and national accountability purposes, but because we know that many programs are planning to use the data to look at effectiveness and program improvement, we plan to address this use as well.

Building the Validity Argument for COS Data

The validity of the COS process was studied by collecting evidence for 16 validity claims (Exhibit 1.3). Additional information about the claims including the rationale, the associated hypotheses, and the criteria is presented later in the report. The evidence for these claims is drawn from four separate studies, with some of the claims having evidence from more than one study. The overall methodology for the research is presented in the next section; then follow four sections, one on each of the studies, presenting additional methodological information and findings. Section 7 contains synthesis of the evidence for each of the claims. The final section contains a review and discussion of the overall findings for the research along with a summary of the accomplishments of the project with regard to its three objectives.

Exhibit 1.3   Validity Claims for the COS Process

A	The three outcome areas reflect important outcomes for accountability and program improvement.
B	The skills and behaviors described in the COS guidance materials and training resources are sufficient to enable providers to accurately describe the content of the three OSEP outcomes.
C	There is variability in children's functioning in the three outcome areas, and that variability is reflected in the COS ratings.
D	There are developmental sequences within each outcome that provide the internal structure of the COS ratings.
E	Providers of EI and ECSE services can be trained to understand and correctly apply knowledge of child development and the COS rating criteria such that a child's team will assign an accurate rating.
F	Functioning in one outcome area is related to functioning in another outcome area.
G	COS ratings in the corresponding outcomes are moderately related with the social emotional (outcome 1), cognitive (outcome 2), communication (outcome 2), and adaptive (outcome 3) domain scores of assessment tools.
H	COS ratings will be related to the type and severity of the child's disability.
I	COS ratings will not be related to the composition of the team, the particular assessment tool used, or child characteristics such as gender or race/ethnicity among children with the same types and severities of disabilities.
J	COS entry scores will be similar for programs and regions serving similar populations.
K	COS rating distributions at entry will be related to the disability-related characteristics of the population served by the states.
L	Similar populations of children enter programs each year, so functional levels reflected in COS ratings should remain constant without intervening factors (e.g., new eligibility criteria, rigorous quality assurance, or improvement process implemented).
M	Functioning, as reflected in the COS rating, in an outcome area at time 1 is related to functioning in that outcome area at a later point in time.
N	The rating structure of the COS is sensitive to both improvements in and maintenance of developmental trajectories that occur in effective programs (i.e., COS ratings differentiate effective from ineffective programs).
O	Data produced by the COS are sufficiently precise to enable states to track the overall status of their EI or ECSE system with the summary statements and monitor change toward targets on those summary statements.
P	Providers will report minimal negative consequences in practice as a result of implementing the COS process.

Section 2 — Overview of Project Methods

Overarching Project Objectives

As described earlier, the COS is in widespread use nationally with children birth through 5 years who have disabilities or other special needs in part because it supports multidisciplinary best practice in early childhood assessment and is consistent with the approach promoted by numerous professional organizations. To date, the COS process has not undergone psychometric validation. Given its widespread use, there is an urgent need to understand the conditions under which the COS process produces valid and reliable data for early intervention and early childhood special education accountability purposes. This project called ENHANCE[5] addresses three objectives related to the validity of COS data: (1) conduct a program of research to examine the validity of ratings generated by the COS process and identify conditions that lessen validity, (2) revise the COS form and supporting materials based on study findings, and (3) identify a series of validity analyses that can feasibly be conducted in states to allow each state to examine the validity of its own COS data on an ongoing basis.

Overview of Studies

As described, the validity argument (Kane, 2006) for the COS consists of 16 claims (Exhibit 3) related to the content, response process, internal structure, relationship to other variables, and consequences of use of the COS (National Research Council, 2008). Drawing on the guidance of national experts in measurement and statistics, we conducted a series of four studies to examine these claims.[6] Taken together, the four studies examined the content of the COS, the rating process, the constructs addressed, the relationship of COS ratings to other variables, and the consequences associated with COS use.

- **Studies 1 through 3** were carried out in 35 local programs, including 19 early intervention and 16 early childhood special education programs in eight states.
  - **Study 1 (provider survey)** examined team members' perceptions of facets of COS use such as perceived impact on practice, provider knowledge of the outcomes, and their self-assessment of the degree of difficulty in applying the criteria ($n = 856$ providers, including 472 EI providers, 302 ECSE providers, and 82 who worked in both EI and ECSE). Specifically, this study used online surveys of providers in

---

5 This IES-funded project was called ENHANCE to use as its name with participating states and programs, with its focus on understanding how to enhance the quality of child outcomes and the objective of EI and ECSE programs is to enhance child outcomes.

6 A fifth proposed study was dropped because we discovered that there was relatively poor documentation in local program records and the limited resources of the grant were better spent on the other more promising studies. The proposed study was intended to examine fidelity of implementation and develop procedures for state use through record reviews of completed COS forms.

programs using the COS to learn about their impressions of the COS, the process, the training and support they received, and the impact of the COS on their practice.

- o **Study 2 (comparison with child assessments study)** examined the relationship between COS ratings and two standardized child assessment tools, the Battelle Developmental Inventory, Second Edition (BDI-2) and the Vineland Adaptive Behavior Scales, Second Edition (Vineland-II) at two time points in 35 programs in eight states. The study examined relationships among these three approaches as well as the differences in the three approaches with regard to producing data for OSEP (n = 153 children with entry data and n = 70 with exit data). Specifically, this study involved comparing the COS ratings with two other commonly administered tools (the BDI-2 and Vineland-II) to learn the extent to which children were classified in the same progress category by each of these approaches. The analysis therefore examined how both the agreement on progress categories and the extent to which COS ratings were related to scores on the assessments at a single time as a secondary purpose of this study.

- o **Study 3 (team decision-making study)** examined fidelity of COS implementation, including accuracy of the ratings through analysis of videos of teams making decisions about COS ratings (n = 113, comprising 63 EI teams and 50 ECSE teams). Specifically, in this study video recordings of actual team implementation of the COS decision-making process were collected to examine both structural and process elements associated with identifying COS ratings and completing the form. The videos were coded to determine whether or not the team reached a rating that was consistent with the available evidence presented and examined team members' knowledge and application of knowledge of the content of the three outcomes, the rating criteria for each of the 7 points on the rating scale, and accurate use of child development behaviors and age expectations.

- **Study 4 (extant state data study)** examined multiple validity claims using extant data by analyzing datasets from 18 state programs (9 Part C, 9 Part B Preschool) using the COS. Specifically, in this study, extant data from state programs and national data compiled by the Early Childhood Outcomes center and the Early Childhood Technical Assistance Center using the COS were analyzed to examine specific claims shown Exhibit 1.4. Examples include the distributions of COS ratings, relationships across outcomes, relationships between time 1 and time 2 ratings, and relationships to other factors such as the child's type of disability.

On the basis of the findings from these studies, we expected that the resulting data would have implications for making any necessary revisions to the COS and the supporting materials such as the instructions and the training materials (see Section 8 of this report for a description of these implications).

The section below contains general information about the four studies and their samples and timelines for their implementation. Additional details about the methods for each study and the validity claims tested in each study are in the separate chapters that describe each of the studies.

General Project Information

ENHANCE studies focus specifically on the data being collected through the use of a process that involves team decision-making and multiple sources of information to arrive a rating of the child's current functioning. As described in the Background, the Child Outcomes Summary (COS) process involves a team of people who know the child discussing the child's functioning and reaching consensus on an appropriate ranting in each of three outcome areas. These team ratings of child functioning in the three outcome areas are then recorded on a Child Outcomes Summary Form (COSF[7]). Because programs report these data to states and states report the data to the federal government, it is essential that the data being produced through this process be valid and reliable.

Selection of State Agency Sample

Investigation of many of the validity claims required the collection of new data from local programs within states. The research team first recruited states for participation in the ENHANCE project and then worked with the state agencies to recruit local programs or school districts for participation in Studies 1, 2, and 3. A second round of recruitment of states to participate in the extant data study (Study 4) was conducted and is described in detail in the methods section in the chapter on that project (Section 6).

An orientation to the ENHANCE project was provided at the annual *Measuring and Improving Child and Family Outcomes Conference* held for and attended by many Part C (EI) and Part B Preschool (ECSE) state staff. Information also was shared via webinars, communities of practice, and individual conversations with states. Details were disseminated about what participation involved, and states were encouraged to participate in ENHANCE. In the end, seven state EI agencies and six state ECSE agencies located in eight different states agreed to participate in the project and help identify local programs to participate in Studies 1 through 3. We attempted to select the state agency sample so that both the EI and ECSE agency from a given state were participating to contain costs of data collection at the local level. The state agency sample was selected to include states that varied with regard to population size and region of the country, with sampled states geographically distributed. A final set of eight participating states were recruited: Illinois, Maine, Minnesota, New Mexico, North Carolina, South Carolina, Texas, and Virginia (see Exhibit 2.1). Of those eight states, five included both EI and ECSE, two included EI only, and one included ECSE only.

---

[7]   When initially implemented and when this grant proposal was written, the COS process was referred to by many states as synonymous with the COSF. Since then, language has shifted to refer to the process as the COS and only speak about the COSF in direct reference to documentation or other issues with the form itself.

Exhibit 2.1   States Participating in ENHANCE Studies, by EI, ECSE, or Both EI and ESCE

Legend
EI and ECSE – Bright Green
EI Only - Yellow
ECSE Only - Blue
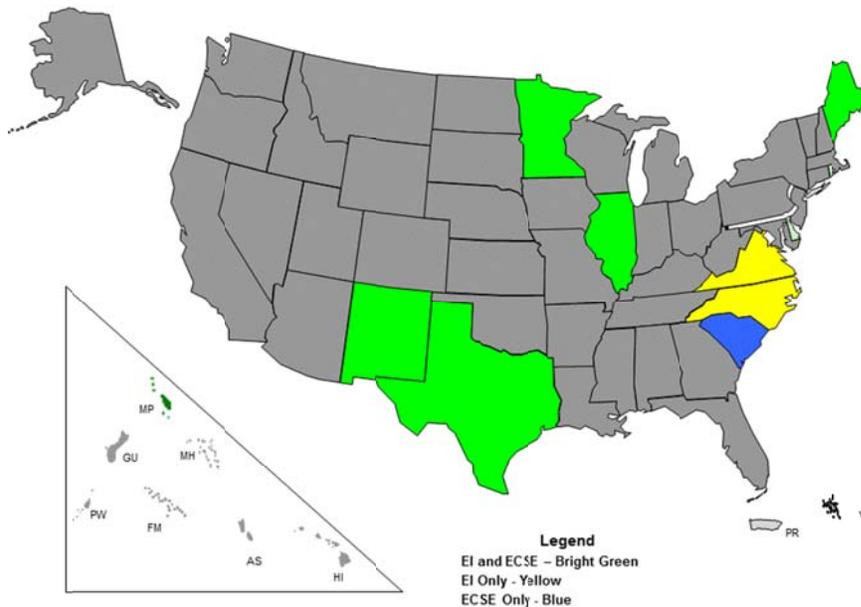
Selection of Local Areas and Programs (Studies 1 through 3)

Studies 1 through 3 required data collection from a sample of local programs within the states participating in ENHANCE (Exhibit 2.1). Participation of EI and ECSE local programs was necessary to examine a range of validity claims. Local programs provided samples of children, team members, or entire teams for each study, with specific criteria depending on the study and described in later chapters.

In each state program, we attempted to select three local programs, with an effort to identify EI and ECSE programs from the same local areas whenever possible to reduce data collection costs in states where both EI and ECSE programs participated in the project. We worked closely with the state to identify potential local areas in which to conduct the study. Our goal was to identify local areas of different sizes and from different geographic regions of each state. We sought participation from local programs that were interested in participating in the study, were implementing the COS process well, and were in a position to make a commitment to participate throughout the 4-year study. Our intention was to select local programs that volunteered to participate and were implementing the COS process well (e.g., received staff training and feedback about the COS process consistent with best practice in the state; were consistently completing COS forms) in order to test validity claims under conditions of good implementation.

For ECSE, the local administrative body was a school district or a regional entity that administered ECSE (e.g., a regional agency that serves several school districts). For sample purposes, we considered the school district or the regional agency to be the local program. In large urban areas, we selected a smaller geographic area within the district for the study, because

we did not need to include the entire district. In one instance, two districts that received referrals from the same referral center and eligibility assessment unit in the area both participated so we could access the full distribution of children needed for studies. For EI programs, the local administrative body varied, as did its boundaries, and was consistent with however the local area was defined by the state. Throughout the rest of this report we refer to the local administrative body and area in EI programs and local school districts as local programs.

ENHANCE staff encouraged states to help us identify local programs with the following characteristics: (a) serving children with a variety of disabilities, (b) serving more than 100 children, (c) implementing the COS process well, and (d) able to participate for the full duration of the project. If possible, in states with both EI and ECSE participation, project staff tried to identify sites that were geographically close so that assessors needed for the child assessment study (Study 2) could assess children in multiple sites. Throughout the project, some local programs had difficulty participating, declined to participate, or in one case was dropped because no assessors for Study 2 could be found within 150-mile radius. When this occurred, we worked with the state agency staff to identify a replacement program in the state, or if that was not possible a replacement local area was identified from another state already participating in the project.

The research team worked with the state agency staff to identify potential local programs and to recruit them into the study. State agencies varied in their approaches to identifying local programs. Some requested programs to volunteer, others spoke individually with directors in programs that were the best candidates, and still others identified programs or districts to include and informed the sites that they had been selected to participate unless the program voiced concerns. Once the local programs were selected and had tentatively agreed to participate, a member of the research team worked with providers and other program staff (e.g., program directors) in the sampled local programs in year 1 to acquaint them with the goals of the studies, the types of information that would be collected, and how families would be recruited. If desired, ENHANCE staff provided orientations via webinar or phone to staff in local programs as well. Local programs received materials providing specific information about each of the studies as well as details about procedures and materials needed to recruit study samples. We asked local staff to obtain parent permission for Studies 1 through 3 because those studies required access to personally-identifiable child-level information. ENHANCE project staff scheduled regular phone contact or email check-ins with programs throughout the project to communicate about ongoing recruitment, the status of children/families in the sample, the status of paperwork, and upcoming activities for the project.

The final local program sample was 19 EI programs and16 ECSE programs across the eight states. For six states, both EI and ECSE local programs in the state participated (15 EI and 13 ECSE local programs), while for two states only EI programs participated (4 local EI programs) and for one state only ECSE programs participated (3 ECSE programs).

Project Timeline and Local Child and Provider Samples (Studies 1–3)

Studies 1–3 were implemented in the same local programs. However, these three studies did not draw on identical samples of children and provider teams (see additional information about characteristics of the samples for Studies 1–3 in Sections 3–5 below). Thus, recruitment of participants, data collection, and analysis were conducted based on appropriate timelines for each study. Information was gathered about the extent to which samples from these studies overlapped so that information was available during analysis.

Data collection for Study 1 (provider survey) took place in the third year of the project. This study involved a total sample of 856 providers; this included a total of 554 EI providers and a total of 384 ECSE providers, of which 82 providers delivered services for both EI and ECSE. Data collection for Study 2 (comparison with child assessments study) began in the first year of the project to allow children who participated in programs for 3 years to contribute both entry and exit data to the project. The final sample for this study was 153 children, with 70 children who had both entry and exit data and were included in longitudinal analyses. Data collection for Study 3 (team decision-making study) took place across the third and fourth years of the project. The final sample for this study was 113 children for whom videos of team decision-making with the COS were included in the final analysis.

Challenges and Contextual Features

Implementation of the four ENHANCE studies as originally designed proved to be difficult due to a number of challenges and state and local contextual features. Some challenges were the result of the specific time period in which the project occurred while other challenges related to state and local issues facing EI and ECSE systems and programs more generally.

Across all four planned studies, recruitment of states and local programs for the ENHANCE studies beginning in 2010 was challenging because of the great recession. Many state EI and ECSE programs were confronted with significant budget cuts, which often translated into layoffs of state or program staff, reductions in funds for their service systems, changes in eligibility requirements and/or introduction of cost sharing with family fees for EI and/or ECSE. In some cases, reorganization of local programs or service regions meant that local program staff served more families across larger geographic regions. At the local level, all of these developments meant that local programs were being asked to do more with fewer resources. All these changes made it difficult for state agencies and local programs to commit to and participate fully in a 4-year project. In particular, these contextual factors impacted the project in the following ways.

- To address budget cuts, some state agencies had cuts in state staff which made recruitment of states as well as local programs difficult. It also meant that states had less staff available or less experienced staff for seeking approvals for data sharing agreements, exporting data, or conducting basic analyses as required in the extant state data study (Study 4).

- Recruitment of local programs also was challenging in this climate of fiscal uncertainty and staff turnover. In some cases, state agencies were unable to solicit the participation of those local programs with the best COS implementation. In these cases, state staff determined which local programs they could expect to participate or who would be most likely to want to participate. These approaches may not have resulted in the selection of local programs with high commitment to participating in the project or who had high-quality COS implementation.

- Increased staff turnover at both state and local levels meant that , commitment to participation in the ENHANCE project fluctuated over the 4-year study period and/or local staff had limited or delayed follow-through to help recruit families or submit required data in a timely fashion. Turnover required greater communication about the project, its goals, and data collection requirements as these had to be continually explained as new staff were hired over time. Lags in replacement of staff sometimes resulted in missed opportunities for follow up with families already enrolled in the study as well. In some cases, staff turnover involved a complete elimination of a position (e.g., preschool special education coordinator) in a whole district or region and the new liaison to the study was unable to fully implement both old and new job responsibilities for day-to-day activities and did not sustain timely support to study activities.

- Data collection at local programs also was impacted adversely because many program staff felt overwhelmed due to budget cuts that forced staff to take on multiple roles and larger caseloads.

Another major challenge was that the COS process had not been introduced in most states for more than a few years when the project began. There were implications for the planned studies at both the state and local levels.

- Many states had recently completed the shift from collecting COS data in a few sites in their states to scaling up statewide. States had to provide training to their own staff and contracted providers. States also began to require data collection of COS ratings and develop systems to monitor the completeness and quality of resulting data. All of this was being done without additional funding. This state of affairs meant that the conditions under which the ENHANCE studies were implemented were less than optimal for testing the validity of the COS process (i.e., examining validity is best undertaken with good or optimal implementation conditions).

- This new need for states to collect and report statewide child outcomes data as children enter and exit EI and ECSE programs highlighted for many states the fact that their existing data systems were not of high quality. As one example of its impact on the planned studies, the research team discovered that local programs and also state agencies did not have accurate and/or timely data about children's exit status, making it hard to complete exit data collection for Study 2 and 3. This also made it difficult to know the extent of missing exit data in the extant state data study (Study 4).

- Local programs and staff varied in their level of commitment to and in some cases understanding of this relatively new COS process. Some programs were more invested in participation than others which the research team felt had impacts on the data collection as well as recruitment of families.

o Data to be presented below from the provider survey (Study 1) support this speculation about staff commitment. For instance, findings from the provider survey indicated that many providers cannot articulate why they are collecting COS data or what happens to the data, and many of them were neither very positive nor very negative about their participation in the COS process.

o Other provider survey data indicated that many providers did not see how the COS process is valuable or has any benefit for families they serve, which may have led providers to view their participation in the project as an added task with little benefit. The research team experienced data collection challenges around maintaining regular communication with local staff about recruitment and timely data collection as well as sharing and sending required information to the research team which seemed to be related to this level of commitment of local programs and staff (e.g., programs submitted forms many weeks after a COS meeting had occurred, impacting available time for conducting child assessments or staff did not notify the research team when children were exiting the program resulting in delays or missed opportunities for exit data collection).

o A related issue about commitment and participation in the project existed for states that use local contracted staff to provide EI or ECSE services. In such states, local program staff told the research team that these contracted staff had little incentive to participate in the project mainly because they could not be reimbursed for the additional time needed to participate in the project.

o Turnover of local program staff often led to delays in implementing data collection, but also meant that the research team had to orient new staff and gain their commitment to working with the research team on the project. Some of these newly hired local staff were hesitant to participate in the project while they were adjusting to their new job.

o In a few cases, entire programs dropped out of the project after participating for some time. Although replacement programs were recruited into the project as soon as possible, this situation led to a delayed implementation of the data collection and resulted in missing data for Study 2 that needed both pre- and post-assessment data) for all children in the study.

o In a number of cases, local program staff were uncomfortable talking with families about the ongoing COS process, especially when families were not present for the team meetings where ratings were decided. For instance, some programs did not want to recruit families for the team decision-making study because getting consent required explaining about the presence of these meetings to families. In other places, discomfort discussing the process with families was a factor influencing the efforts and success of specific providers to enroll families into project studies. A few data collection challenges related to the families themselves and included the following issues.

- Some local programs had difficulty identifying appropriate and sufficient numbers of families that fit the eligibility criteria for the planned studies, including some programs having fewer than expected numbers of families enrolling in EI or ECSE during the study recruitment period, as well as usual attrition and having many transient families who do not participate in EI or ECSE for at least the required 6 months (e.g., migrant or military families) or did not have English as a primary language.

- Program staff reported to the research team that some families were reluctant to participate in the project because either they felt overwhelmed by their child's recent diagnosis, were resistant to having assessments conducted in their home, were reluctant to provide consent for the videotaping portion of Study 2 and/or 3, as well as having difficulties in scheduling the visits to their home (for Study 2).

A final set of challenges in conducting the planned studies concern the limitations in (a) implementing some of the data collection requirements that were originally planned, (b) conducting the data analyses originally planned by the research team, and (c) interpreting the findings. These challenges included the following.

- Because the provider survey (Study 1) indicated that many providers had very limited training on the COS process and states had only been implementing this data collection for a few years, the research team has limited information to confirm the quality of the COS data across all studies.

- None of the samples in any of the four studies are nationally representative samples (and were not planned to be). Rather, the samples were intended to be indicative of high-quality COS implementation. Indeed, there was wide variation in the implementation of the COS process across local programs, both within and across states (as seen in the team decision-making study – Study 3).

- For Study 2 and 3, final sample sizes were smaller than planned. Power for initial study designs were based on analyzing combined data for EI and ECSE programs. With smaller numbers, it was not possible to disaggregate data by site or state, and sample sizes were further diminished in any analysis by EI versus ECSE or by other important predictor variables (e.g., age, disability categories, entry versus exit meetings). For Study 4, such disaggregation of state-level data was not possible because states did not have available data on various child or disability categories in state datasets provided to the research team.

- One measurement issue that affected Study 2 and 3 was the limited availability of high-quality reliable measures to distinguish between children with different types of disabilities or to describe the severity of their disabilities and the influence of these disabilities on their functioning. Measurement challenges partly result from variability across states and local areas in how primary disability and eligibility categories are used in EI and ECSE and the overarching issues that these programs serve children whose challenges are highly variable in complexity and severity. While the research team selected the ABILITIES Index to provide additional information and provide some measure of severity, we recognized that it may not fully capture critical distinctions most relevant to functioning in the three outcome areas across all children. For the child assessments study (Study 2), the research team had difficulty identifying experienced assessors, and the rigorous training and certification process implemented to assure their

ability to conduct the assessments in a consistent standard manner took much longer than expected (with some assessors not achieving the level of quality needed). Given the long time period from the entry to the exit assessments, new assessors had to be hired and trained fairly regularly throughout the project. In one region, some programs recruited families to participate with short timelines remaining for assessment, but limitations on the number certified assessors ready in the region meant that they could not complete all the assessments at times that worked with family members' schedules during the remaining assessment window.

- Because of the service system challenges described earlier, in most local programs the COS entry and COS exit teams for Study 3 were comprised of different individuals. This change in team composition made it difficult to draw conclusions about whether differences observed in the data collection from entry to exit result from the timing of the meeting, the available information teams had access to at the time of the meeting, and/or the variable team composition.

- For Study 4, the research team learned from working with state datasets that states had limited ability to know how much missing data they have and the characteristics of those children with missing data.

- For Study 4, data sets provided are typically entry data paired with data from children who have exit COS data used for federal reporting. These paired entry data may differ somewhat from a broad sample of all children entering EI or ESE regardless of available exit COS data. However, lack of information about missing data makes it difficult to know how much impact this distinction in data may have.

- As mentioned above, the team videos for Study 3 were mainly collected in 2011 and 2012, which was early in COS implementation during the recession that adversely affected states' and local programs' capacity for full implementation of statewide COS training and monitoring as well as state and local programs' full participation in the project.

- The length of time for participation in EI and ECSE services varies for different children, especially when children achieve their goals or family situations change and the child exits earlier than expected based on age. The variable timing expected for exits meant that it was difficult for project staff to anticipate when program staff might not be submitting needed exit data for a child.

- A number of children are in EI and ECSE services for 3 years or more between program entry and exit. Delays enrolling children in studies, especially for Study 2, meant that many children had not yet completed services with their local programs when the project data collection ended.

Section 3 — Study 1: Provider Survey–Design, Methods, Key Findings

Study 1 produced qualitative information about the implementation the COS process by EI and ECSE providers. It explored both the positive and negative consequences associated with implementation of the COS process for accountability purposes—another element of validity in large-scale accountability. The purpose of this study was learn about EI and ECSE providers' perceptions about the use of the COS process and its impact on their practice (claims, B, E, I, and P). For Study 1, we conducted an online survey of EI and ECE providers team members about the implementation and consequences of the COS process, especially about how service providers participating in the process perceived its impact on their practice, whether or not they believed teams were inflating exit ratings to reflect favorably on their programs, and whether providers could be trained to associate functional skills and behaviors with the appropriate outcome. We investigated factors that could serve to enhance or lessen the validity of the COS process (e.g., knowledge of the outcomes, the rating scale, age expectations; who was present during COS decision-making). Survey questions examined providers' understanding of critical issues such as why data are being collected, their self-assessment of the degree of difficulty in applying the criteria, and how frequently teams had difficulty in reaching consensus. Data about similarities and difference in responses for EI and ECSE providers and about provider characteristics also were collected (e.g., years of experience, COS process training, and the approximate number of completed COS ratings). This study also yielded information about implications for revising the COS process and for developing more effective guidance related to COS process use.

Research Questions for Study 1

The objectives and associated research questions for Study 1 are shown in Exhibit 3.1. For this study, our main objectives were to determine (1) how well providers understand the COS process, (2) how well they implement it, (3) their perceptions of factors related to its implementation and its effects on their practices, (4) their perceptions of problems in implementing the COS process, and (5) whether perceptions were similar or different for EI versus ECSE providers. An additional objective was to determine implications for modifications to the COS process or 7-point scale and for its implementation

Exhibit 3.1   Objectives and Research Questions for Provider Survey

1. Determine providers' perceptions about the critical aspects of COS process implementation.
    -   Do providers report that they understand why the data on child outcomes are being collected?
    -   Do providers believe that they understand the three outcomes?
    -   Do providers believe that they understand the rating scale?
    -   Do providers believe they have sufficient knowledge of child development to assign COS process rating accurately?
    -   Do providers believe other team members understand the outcomes, understand the rating scale, and have sufficient knowledge of child development to assign ratings accurately?
    -   What do providers see as the impact of having the parent present during the rating determination?
    -   Do providers believe they have received sufficient training to provide accurate ratings?
    -   Do providers believe they are receiving adequate ongoing support to provide accurate ratings?

2. Determine providers' perceptions about the impact of COS process implementation on their practice.
    -   What positive impacts on practice do providers identify that result from COS process implementation?
    -   What negative impacts on practice do providers identify that result from COS process implementation?

3. Determine providers' perceptions of problems related to implementation.
    -   Do providers believe providers are inflating the exit ratings to show the program as effective?
    -   Do providers believe that other providers take the rating process seriously and understand how to make ratings accurately?

4. Examine the relationship between providers' perceptions of the COS process for EI versus ECSE providers.
    -   Do providers' responses vary for EI and ECSE?

5. Identify the implications of what was learned from the survey for modifications to and support for implementation of the COS process (e.g., training, more specification of how to approach the ratings discussion, provision of materials on typical child development).

Sample for Study 1

The local programs were recruited to participate in Study 1 from the 8 participating states. In the end, providers from EI and ECSE programs in 8 states completed the survey.

Provider sample. All providers in local programs and districts participating in ENHANCE were invited to participate in an online survey in spring 2012. Participation was voluntary. We anticipated that an average of 10 providers per program area would yield a sample of 360 providers. The expected response rate was 70%, which would yield 252 provider surveys. We expected the response rate to be high because the research team worked to build strong relationships with the local programs that participated in multiple studies. The final sample of

providers completing the online survey was greater than planned, with a total of 554 EI providers and a total of 384 ECSE providers, of which 82 provided services for both EI and ECSE (total $N = 856$).

Methods for Study 1

Measure

An online survey was used for data collection (see Exhibit 3.2 for key content areas of this survey; a copy of the provider survey is provided in Appendix D). The survey included questions about the providers' understanding of the COS process and related constructs as well as providers' characteristics, such as years of experience and professional role.

Exhibit 3.2   Content of Provider Online Survey

- Perceptions of COS process and experiences:
    - Ease/difficulty of the rating process and team discussions
    - Ease/difficulty applying rating criteria
    - Accuracy of final team ratings
    - Ability of teams to reach consensus
    - Factors interfering with effective COS ratings
- Estimate of average length of COS team meeting
- Self-assessment of knowledge of
    - Age-expected child functioning
    - Distinctions between functional behaviors and discrete skills
    - What is meant by each outcome
    - The rating criteria
    - Why data are being collected and how to explain the reasons to others
- Own confidence in and competence with the rating process
- Perception of colleagues' competence with the rating process
- Influence of COS use on their knowledge, skills, practices
- Suggestions for improving COS process, supporting materials
- Provider information
    - Profession
    - Years of experience working with young children
    - Demographics (age, gender, race, ethnicity),
    - Ages of clients/students served
    - Number of completed COS
- Information about training and support with the COS process
    - Approximate number of hours of and format for training received
    - Perception of adequacy of the training
    - Ongoing feedback related to COS process
    - Extent of feedback
    - Perceived adequacy of feedback

Data Collection Methods

We field-tested the survey with providers in several states to check for clarity, relevance, and potential improvements before finalizing the version to be completed online. The research team collected email addresses from local programs participating in the study for all providers and sent providers an invitation to participate along with a link to the web survey. SRI staff who worked closely with each program's liaison and with other program staff to follow up with providers to increase the response rate. The survey was completed by providers confidentially. Providers who completed the survey were eligible to receive one of three $100 gift certificates in a drawing as an incentive.

Data Analysis for Study 1

Survey data were analyzed to yield frequencies for each of the survey items and cross-tabulations to examine differences in responses across EI and ECSE providers. The statistical significance of differences between survey responses of subpopulations (e.g., EI vs. ECSE providers) was assessed using standard statistical tests (e.g., chi-square tests, t tests). Data are reported for the 856 providers who completed some portion of the survey; 836 participants completed all items on the survey. The remaining 20 providers mostly completed information about their training experience and demographic information, with little specific information about their COS opinions and experiences.

Key Results for Study 1

Of the 856 providers who completed the survey, about half (55%) served infants and toddlers in EI, about one-third (35%) served preschoolers in ECSE, and 10% served children in both age groups (Exhibit 3.3). Similarly, about half of them (50%) were early interventionists or teachers, more than one-third (38%) were therapists, and 12% had other types of roles. A majority of the providers had 6 or more years of experience providing services (75%), with only 9% having 2 years or less of experience. The providers had a range of experience participating in the COS process, with about half (51%) having participated in over 30 ratings, 28% in between 11 and 30 ratings, and about one-fifth (21%) with experience with only 10 or fewer ratings. Most of the providers had some but less than 4 hours of COS training (70%), although 5% had none. Providers reported that about one-third of COS ratings (36%) took between 1 to 15 minutes to complete, about one-third (35%) took 16 to 30 minutes, and almost one-third (29%) took more than 30 minutes to complete.

Exhibit 3.3   Characteristics and Training of Providers in Study 1

Characteristic	*n*	%
Age group served (*n* = 856)		
EI	472	55
ECSE	302	35
Both	82	10
Role (*n* = 830)		
Early interventionist/teacher	412	50
Therapists	315	38
Other	103	12
Years of experience (*n* = 824)		
2 years or less	73	9
3 to 5 years	132	16
6 years or more	619	75
Number of COS ratings made (*n* = 856)		
10 or less	181	21
11 to 30	241	28
31 or more	434	51
Hours of COS training (*n* = 856)		
None	39	5
Some, but less than 4 hours	599	70
5 to 8 hours	137	16
9 hours or more	81	10
Amount of time to complete COS rating (*n* = 856)		
1 to 15 minutes	308	36
16 to 30 minutes	300	35
More than 30 minutes	248	29

Most providers reported receiving COS training, but many of them did not receive ongoing feedback or support in the COS process (Exhibit 3.4).

- Four-fifths of providers (82%) reported that someone in their program was available to provide them with ongoing support if they asked for support, and two-thirds of them (69%) reported that their program trains new staff on the COS process.

- Half the providers (50%) reported that someone from their program provided feedback to those involved in the COS process, but 21% reported that nobody was available to do this, and another 29% did not know whether this practice was in place in their program.

- Fewer of the providers reported that their program checked the COS forms after they were completed to check for accuracy (37%), but half (53%) reported not knowing whether such a practice was in place.

Exhibit 3.4   Providers' Self-Reported Training and Feedback on COS Process Provided by Their Programs (*N* = 836)
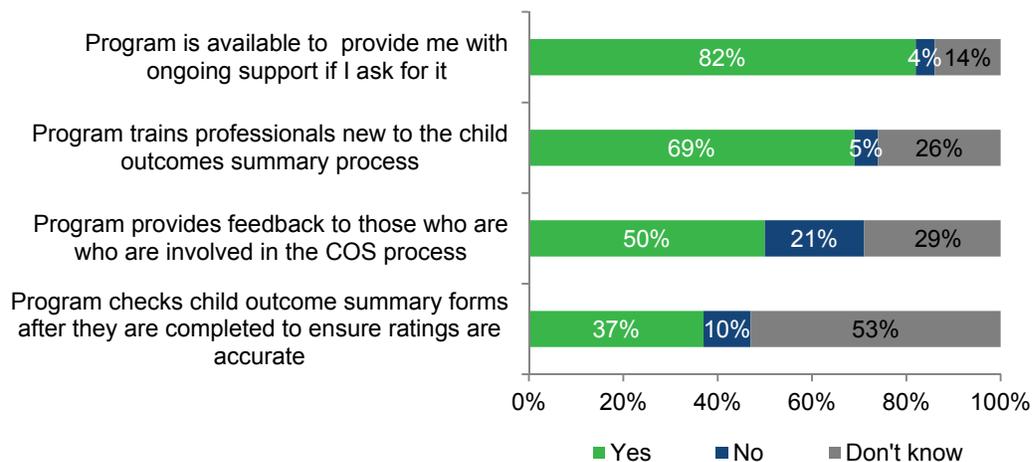
Providers reported variation in the quality of the training and feedback they receive on the COS process (Exhibit 3.5).

- About one-third of providers (32%) reported that that they do not get feedback from someone such as a supervisor on the COS ratings or the form, nor do they get helpful feedback about the COS form (34%).

- Despite those responses, almost half the providers (47%) were in high agreement with the statement that the ongoing support they received about the COS process was adequate, while 44% reported that it was somewhat or a little true for them.

Exhibit 3.5   Providers' Perceptions of Quality of Training and Feedback about COS Process
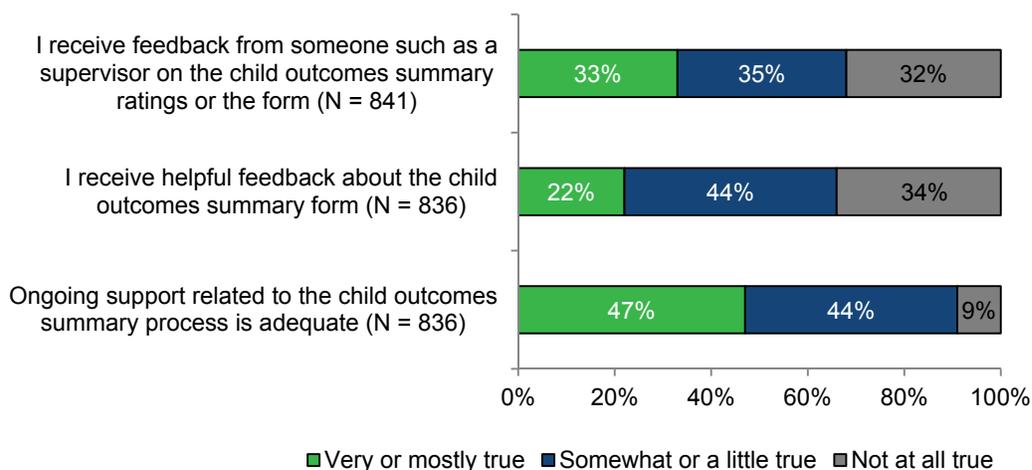
A majority of the providers were confident that they understand the COS process, but they were less knowledgeable about the reasons behind the COS data collection (Exhibit 3.6).

- Nearly three-fourths to almost 90% of providers reported strong endorsement that they understand the meaning of the three outcomes, how to apply the rating criteria, the difference between functional and discrete skills, and what is age-expected functioning in the three outcome areas.

- Many providers were somewhat less aware of the reason they were collecting the child outcomes data (65% reporting strong endorsement) and even less so for knowing why the child outcomes data were being collected (32% reporting strong endorsement).

Exhibit 3.6   Providers' Self-Reported Understanding of COS Process (*N* = 855)
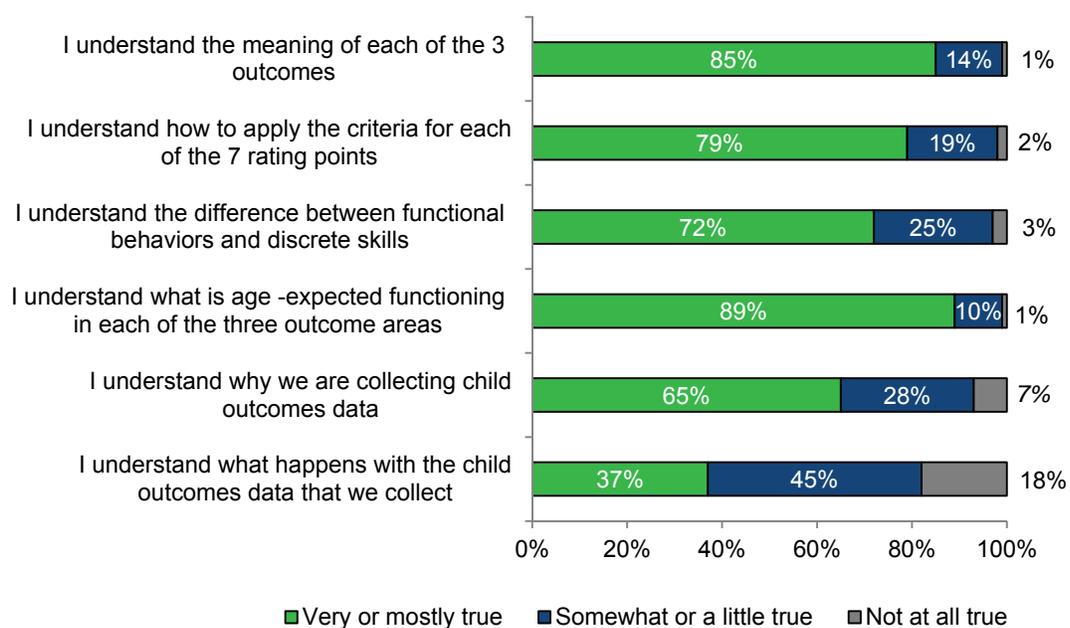
The providers also reported strong confidence that they know how to make the COS ratings, with one exception (Exhibit 3.7).

- Over 80% of the providers strongly endorsed six of the seven statements about their own skills in making the COS ratings.

- In contrast, only a little over half of the providers (56%) strongly endorsed the statement that they know how to explain the need for the child outcomes ratings to families. This finding may not be surprising because many of these providers also reported limited understanding of why the child outcomes data are being collected.

Exhibit 3.7   Providers' Self-Reported Skills in Making COS Rating (*N* = 854)
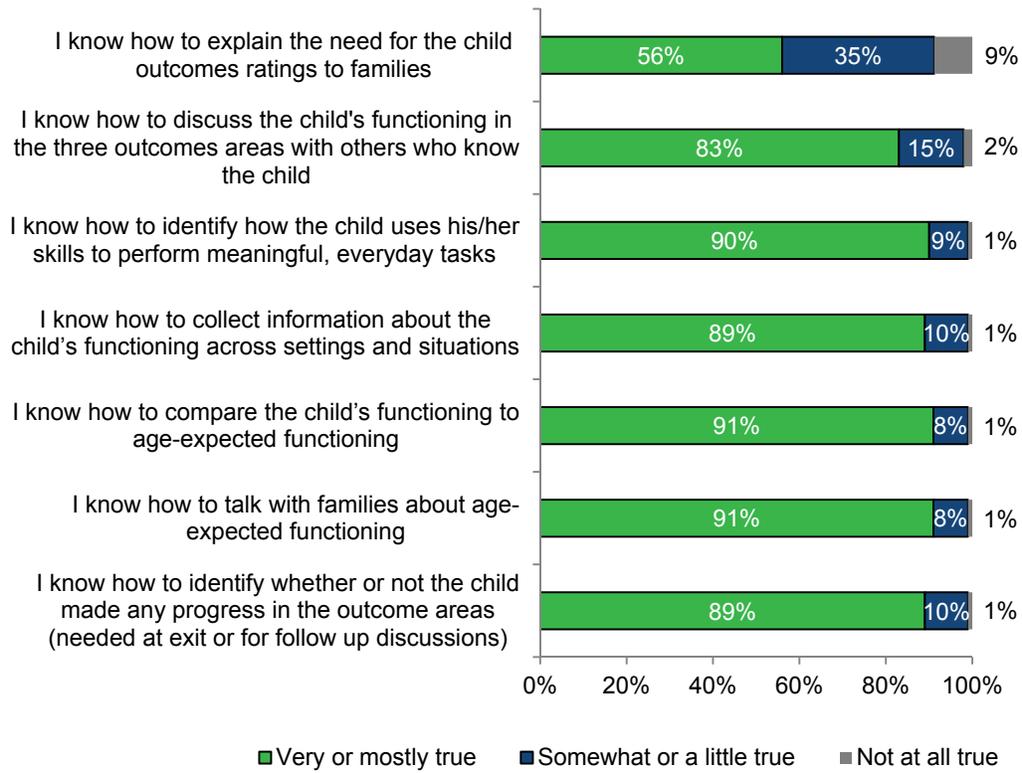
The providers tended to rate their other team members' understanding of the COS process high (Exhibit 3.8).

- However, some providers did report that some, a few, or none of their other team members understand the meaning of the three outcomes (13%), how to apply the rating criteria (14%), the difference between functional behaviors and discrete skills (18%), and what is age-expected functioning in the three outcome areas (9%).

Exhibit 3.8   Providers' Perceptions of Other Team Members' Understanding of COS Process (*N* = 811)
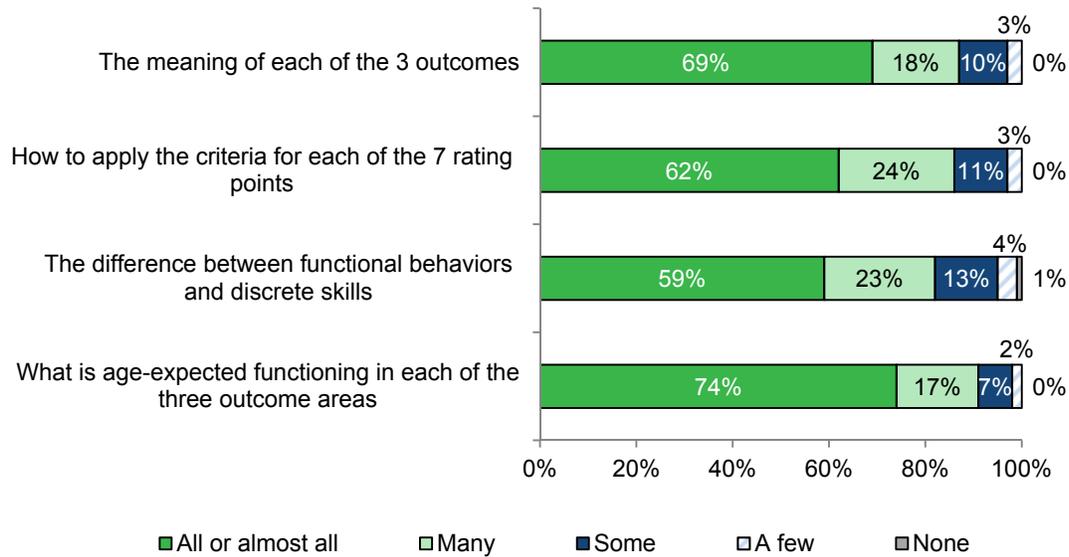
To learn about how the providers actually make the COS ratings, we included on the survey questions about the descriptive features of the COS process and the providers reported on how many COS ratings they had participated in that had each characteristic (Exhibits 3.9, 3.10, and 3.11). Across these ratings, a majority of providers reported that they felt that the COS process was well implemented with integrity.

In terms of who participated in the COS ratings (Exhibit 3.9):

- For most providers (75%), all or almost all the COS ratings were made by a team involving the provider and at least one other professional.
- For about two-thirds of providers (68%), all or almost all the COS ratings were made with the family providing input; but for only one-third of providers (34%), all or almost all the COS ratings were made with the family present when the ratings were made.

**Exhibit 3.9   Providers' Report about Their Experiences Using the COS Process:
Who Participates in COS Ratings (*N* = 849)**
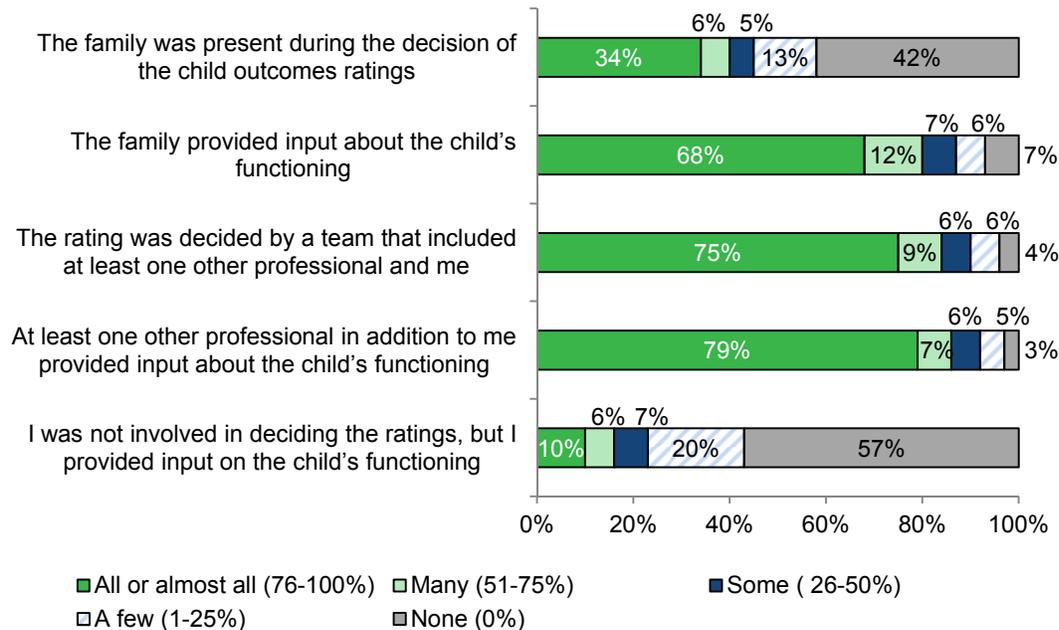
In terms of the types of information used to make the COS ratings (Exhibit 3.10):

- A majority of providers reported that all or almost all the COS ratings were made using information about the child's functioning in multiple settings and situations (75% of providers) and from one or more assessment tools (76%).

- A majority of providers (85%) reported that all or almost all the COS ratings were made with everyone involved considering the information carefully to make the ratings.

- A majority of providers (85%) reported that for all or almost all the COS ratings made there was enough information about the child's functioning in each outcome to make the rating.

Exhibit 3.10 Providers' Report about Their Experiences Using COS Process: Information Used to Make COS Rating
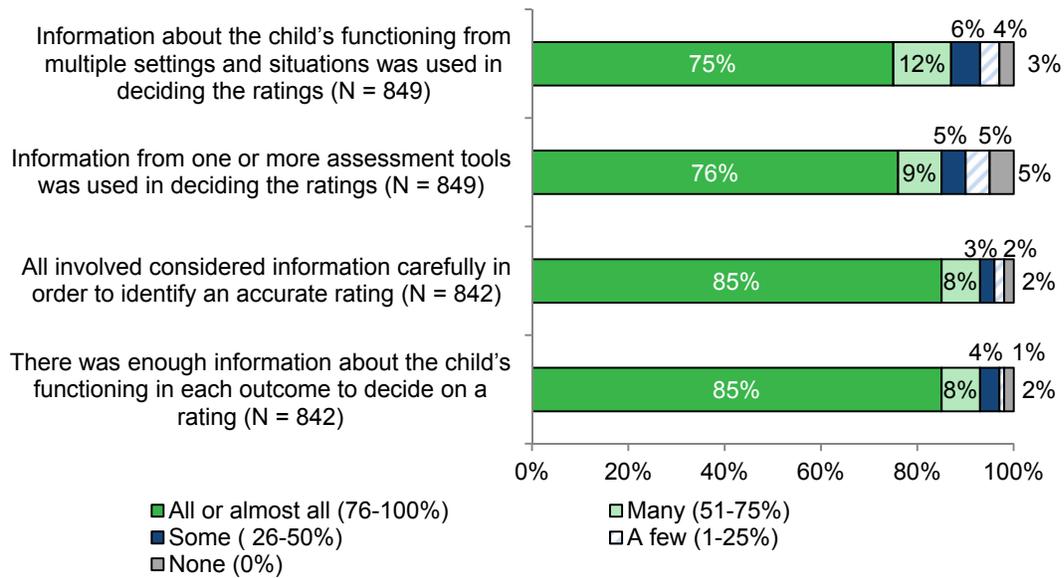
In terms of the providers' perceptions of the integrity of the COS rating process (Exhibit 3.11):

- A majority of providers (80%) reported that for all or almost all the COS ratings there was enough time to review the child' functioning to make the ratings.

- A majority of providers (76%) reported that for all or almost all the COS ratings they were confident that the ratings given were accurate.

- A majority of providers (78%) reported that for all or almost all the COS ratings the rating process matched their understanding of how it was supposed to be done.

- A majority of providers (88%) reported that none of the COS ratings were selected to make the program look good.

Exhibit 3.11 Providers' Report about Their Experiences Using COS Process: Perceptions of Integrity of COS Rating Process *(N = 842)*
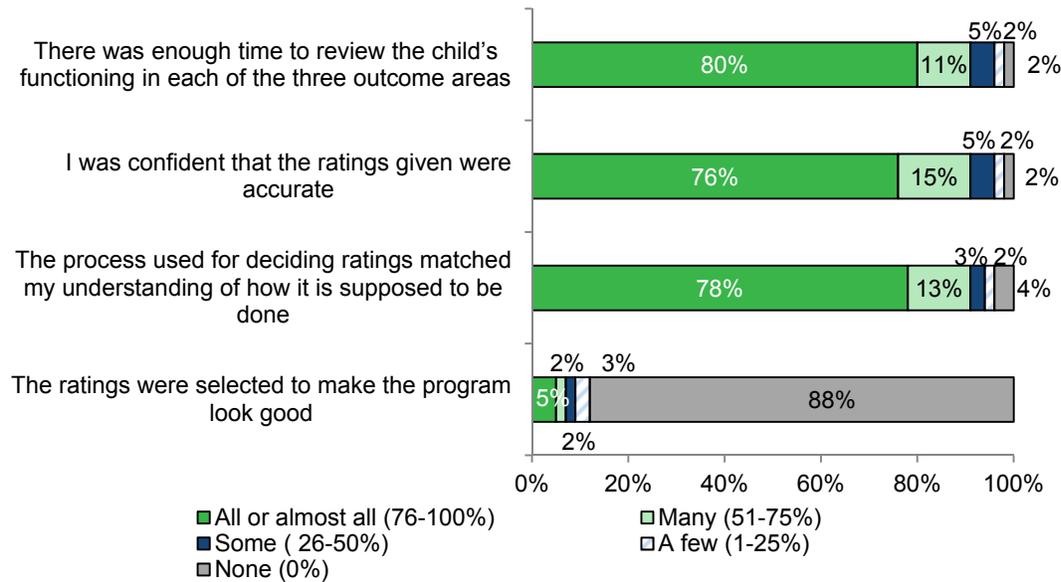
Survey items asking about providers' attitudes about the COS process indicated variation in how they felt about the COS process (Exhibit 3.12):

- A majority of the providers (77%) strongly endorsed the statement that the assessment tools they used were helpful in making the COS ratings.

- Only about one-fourth of providers (26%) strongly endorsed the statement that they liked the three outcomes, and another half of them (52%) reported that this statement was somewhat or a little true for them.

- About one-third of providers (30%) strongly endorsed the statement that the COS process is a good way to collect the child outcomes data, and another half of them (54%) reported that this statement was somewhat or a little true for them.

- About one-third of providers (32%) strongly endorsed the statement that the COS process has too much additional paperwork, and another half of them (49%) reported that this statement was somewhat or a little true for them.

- Only 10% of providers strongly endorsed the statement that the COS process emphasizes age-expected functioning too much, and another half of them (46%) reported that this statement was somewhat or a little true for them.

Exhibit 3.12   Providers' Self-Reported Attitudes about COS Process

	Very or mostly true	Somewhat or a little true	Not at all true
Information from assessment tools we use is very helpful in determining ratings for the three outcomes (N = 841)	77%	21%	2%
I like the three outcomes (N = 836)	26%	52%	22%
The child outcomes summary process is a good way to collect data on child outcomes (N = 840)	30%	54%	16%
There's too much additional paperwork associated with the child outcomes summary process (N = 840)	32%	49%	19%
The child outcomes summary process is a useless activity (N = 836)	21%	51%	28%
The child outcomes summary process emphasizes age-expected functioning too much (N = 836)	10%	46%	44%

■ Very or mostly true   ■ Somewhat or a little true   ■ Not at all true

Survey items asking about providers' perceptions of the accuracy of the ratings derived from the COS process indicated that most felt that the COS process was yielding accurate ratings of children's functioning with little difficulty in reaching consensus on ratings, although there was variation across providers for many of the ratings (Exhibit 3.13).

- About one-third of providers (33%) strongly endorsed the statement that the COS ratings are too subjective, and another half of them (55%) reported that this statement was somewhat or a little true for them.

- Almost one-fifth of providers (18%) strongly endorsed the statement that the COS ratings tend to be low at entry relative to the child's level of functioning , and another one-third of them (37%) reported that this statement was somewhat or a little true for them.

- Only 9% of providers strongly endorsed the statement that the COS ratings tend to be high at exit relative to the child's level of functioning, and another 29% reported that this statement was somewhat or a little true for them.

- Very few providers strongly endorsed the statement that COS ratings are higher or lower than the child's actual level of functioning (3% and 2%, respectively), and another 29% reported that these statements were somewhat or a little true for them.

- More providers strongly endorsed the statement that COS ratings are more accurate when parents are present (22%) than they strongly endorsed the statement that COS ratings are less accurate when parents are present (9%).

- Few providers (5%) strongly endorsed the statement that it is difficult to reach consensus on one of more of the COS ratings.

Exhibit 3.13

	Very or mostly true	Somewhat or a little true	Not at all true
Child outcomes summary ratings are too subjective (N = 840)	33%	55%	12%
Ratings tend to be low at entry relative to the child's actual level of functioning (N = 840)	18%	37%	45%
The ratings given are higher than the child's actual level of functioning (N = 840)	3%	29%	68%
Ratings tend to be high at exit relative to the child's actual level of functioning (N = 836)	9%	29%	62%
The ratings given are lower than the child's actual level of functioning (N = 836)	2%	29%	69%
Ratings are more accurate when parents are present for the rating decision (N = 836)	22%	40%	38%
Ratings are less accurate when parents are present for the rating decision (N = 836)	9%	34%	57%
It is difficult for individuals involved in identifying child outcomes ratings to reach consensus on one or more of the three outcomes (N = 836)	5%	46%	49%

Survey items asking about providers' perceptions of the positive impacts of the COS rating process on their practice indicated variation in positive impacts, with more than half the providers endorsing all impacts as having at least some positive impact on their practice (Exhibit 3.14).

- About one-third or more of providers strongly endorsed statements that the COS process helps them think about children's progress over time (40%), helps them think about children's functioning across settings and people (37%), makes them think about

children's functioning relative to age expectations (35%), helps them focus on functional skills (32%), and helps them discuss the "whole child" (30%).

- About one-fifth strongly endorsed statements that the COS process improves their conversations with families about their child (21%), improves the way they work as a team (20%), improves the assessment process (17%), and leads to better Individualized Family Service Plan (IFSP) or Individual Education Program (IEP) outcomes (16%).
- The vast majority of providers (91%) indicated that the COS process had a positive impact for at least one of the nine items in Exhibit 3.14.

Exhibit 3.14   Providers' Perceptions about Positive Impacts of COS Process on Their Practice (*N* = 831)
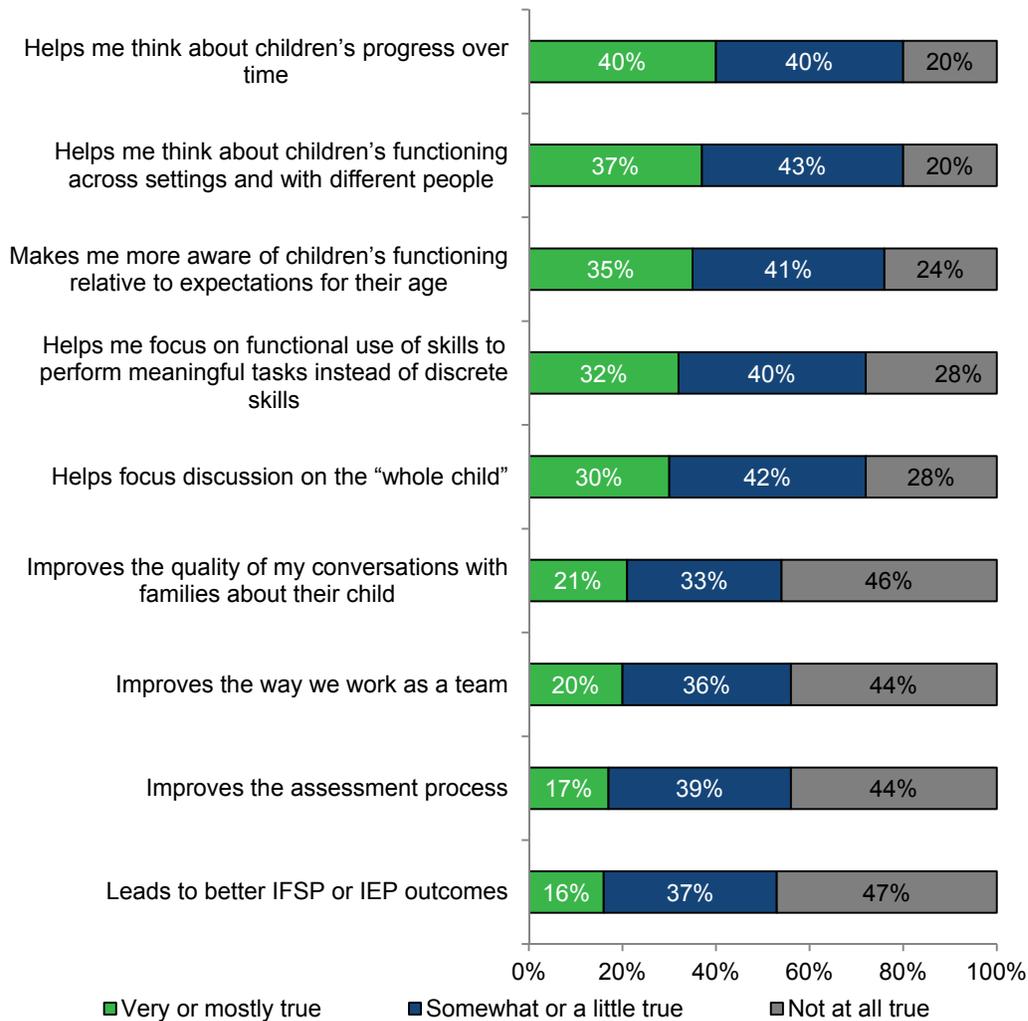
Survey items asking about providers' perceptions of the negative impacts of the COS rating process on their practice indicated very low endorsement of statements about negative impacts, with one exception (Exhibit 3.15).

- About one-third of the providers (31%) strongly endorsed the statement that the COS process takes time away from other important activities, and another 42% reported that this statement was somewhat or a little true for them.
- Only a very few providers strongly endorsed statements that the COS process negatively impacts the assessment process (3%), has negative effects on their relationships with families (2%), or leads to poorer quality IFSP or IEP outcomes (2%).
- A majority of providers (77%) indicated no negative impacts for any of the four items in Exhibit 3.15.

Exhibit 3.15   Providers' Perceptions about Negative Impacts of COS Process on Their Practice (*N* = 831)
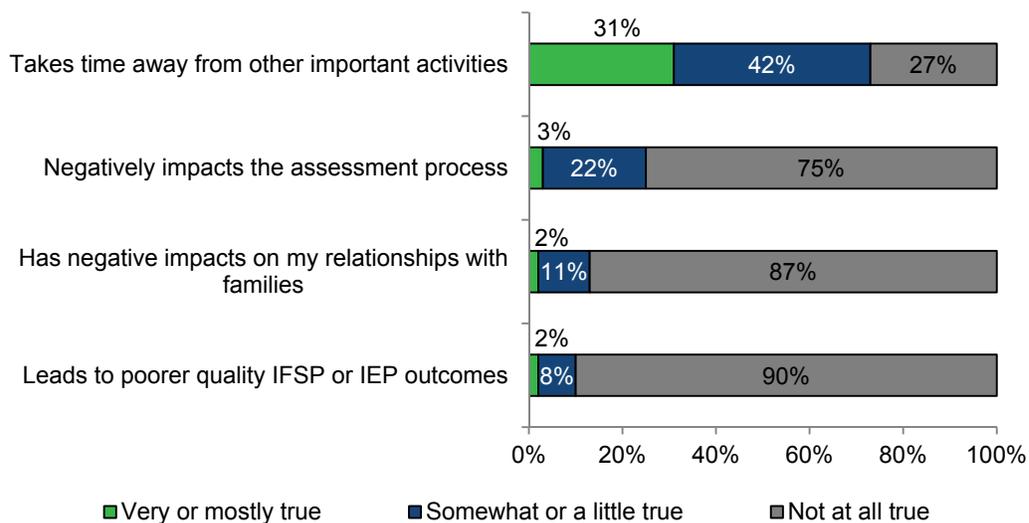
When asked to indicate the overall impact of the COS process on their work with children and families, only 7% identified the process as having a negative or very negative impact on their work. Most (68%) providers felt the COS process had a neutral impact on their work. The remaining 25 percent responded that the COS process had a positive or very positive impact on their work.

Comparison of Results for Study I for EI and ECSE Providers

The results described above for the overall sample of providers were very similar for EI providers and ECSE providers with some exceptions (see complete set of comparable tables for EI and ECSE providers in Appendix E).

- With regard to experience in the COS process, more EI providers have participated in more COS processes than ESCE providers (64% of EI providers have participated in more than 30 COS versus 32% of ECSE providers; *p* < .001) (overall data in Exhibit 3.3).
- With regard to COS training, 52% of EI providers reported having limited COS training (2 hours or less) compared with 33% of ECSE providers (*p* < .001) (overall data in Exhibit 3.3).

- With regard to COS training and feedback on the COS process, there were differences between reports by EI and ECSE providers for three of seven ratings (overall data in Exhibits 3.4 and 3.5).
  - Fewer EI providers (66%) than ECSE providers (71%) reported that their program trained professionals new to the COS process ($p < .01$).
  - Fewer EI providers (34%) than ECSE providers (41%) reported that their program checked the COS forms after ratings were completed ($p < .04$).
  - Fewer EI providers (29%) than ECSE providers (40%) reported that they received feedback from someone like a supervisor on their ratings or the COS from their program compared ($p < .04$).
- EI and ECSE providers reported similarly about their understanding of the COS process with one difference out of six ratings (overall data in Exhibit 3.6).
  - Fewer EI providers (78%) reported strongly agreeing that they understand how to apply the 7-point COS rating scale compared with ECSE providers (82%) ($p < .01$).
- EI and ECSE providers reported similarly about their perceptions of their other team members' understanding of the COS process with one difference out of four ratings (overall data in Exhibit 3.8).
  - More EI providers (77%) reported that many to all of their other team members know what is age-expected functioning in the three child outcomes compared with ECSE providers (68%) ($p < .001$).
- EI and ECSE providers' reports about who was present during the COS process differed on three of five ratings (overall data in Exhibit 3.9).
  - Of the EI providers, 48% reported that the family was present for more than 50% of COS ratings compared with 33% of ECSE providers ($p < .0001$).
  - In 82% of COS for EI, at least two providers gave input about the child's functioning compared with 73% of COS for ECSE ($p < .004$).
  - More of the EI providers (48%) than ECSE providers (35%) reported that there were at least some COS where they were not involved in deciding the ratings but they provide input about the child's functioning ($p < .001$).
- EI and ECSE providers' reports about the information used to make the COS ratings differed on one of four ratings (overall data in Exhibit 3.10).
  - Fewer EI providers (83%) than ECSE providers (87%) reported that all or almost all the COS ratings were made with all or almost all involved considering information carefully in order to make an accurate rating ($p < .02$).
- EI and ECSE providers' attitudes about how they felt about the COS process differed on one of six ratings (overall data in Exhibit 3.12).
  - More ECSE providers (39%) than EI providers (27%) strongly agreed with the statement that the COS process had too much additional paperwork ($p < .001$).
- EI and ECSE providers' perceptions about the quality of the COS process differed on two of eight ratings (overall data in Exhibit 3.13).

- o More EI providers (26%) than ECSE providers (17%) strongly agreed with the statement that the COS ratings were more accurate when parents were present for the rating ($p < .003$).
  - o More EI providers (36%) than EI providers (29%) strongly agreed with the statement that the COS ratings were too subjective ($p < .03$).
- EI and ECSE providers' perceptions about the positive impacts of the COS process on their practice differed on only one of nine ratings (overall data in Exhibit 3.14).
  - o More EI providers (18%) than EI providers (15%) strongly agreed with the statement that the COS leads to better IFSP or IEP outcomes ($p < .01$).
- EI and ECSE providers' perceptions about the negative impacts of the COS process on their practice differed on three of four ratings (overall data in Exhibit 3.15).
  - o Fewer EI providers (29%) than ECSE providers (22%) agreed with the statement that the COS negatively impacts the assessment process ($p < .02$).
  - o Fewer EI providers (17%) than ECSE providers (8%) agreed with the statement that the COS negatively impacts their relationships with families ($p < .00003$).
  - o Fewer EI providers (25%) than ECSE providers (39%) agreed with the statement that the COS takes time away from other important activities ($p < .001$).

Summary, Implications, and Limitations of Study 1

The data from the provider survey completed by 856 EI and ECSE providers across eight states provide many important findings that suggest that the COS process can be well implemented, with both EI and ECSE providers generally reporting that they understand the COS process and that the ways it is being implemented in their local programs is adequate to yield accurate data about children's functioning.

- Most providers reported receiving COS training, but many had very few total hours of training (70% with some but less than 4 hours and 5% with none), and many reported not receiving ongoing feedback or support on the COS process.

- Providers reported variations in the quality of the training and feedback they receive on the COS process, but nearly half of them (47%) reported that they strongly felt the training and feedback they received were nevertheless sufficient to make an accurate rating of children's functioning

- Despite what may seem like limited training on the COS process, a majority of the providers were confident that they understand the COS process and how to make the ratings, although many of them were less knowledgeable about the reasons behind the COS data collection.
  - o Nearly three-fourths to almost 90% of providers reported strong endorsement that they understand the meaning of the three outcomes, how to apply the rating criteria, the difference between functional and discrete skills, and what is age-expected functioning in the three outcome areas.

- A majority of providers also reported strong confidence that they know how to make the COS ratings, with more than 80% of the providers strongly endorsing six of the seven statements about their own skills in making the COS ratings.
- Only about two-thirds of providers (65%) reported strong endorsement that they understand why they are collecting the child outcomes data, with only about one-third (37%) reporting strong endorsement that they understand what happens with the child outcome data that they collect

- The providers also tended to rate their other team members' understanding of the COS process high.

- Across a range of survey items asking about providers' perceptions of the accuracy of the ratings derived from the COS process, the data indicated that most providers felt that the COS process was yielding accurate ratings of children's functioning with little difficulty in reaching consensus on ratings. That is, most providers felt the COS process was being well implemented with integrity.

- Most providers also reported that the COS process generally was not having negative impacts on their practice or on their relationships with the families they work with. The negative impacts with strongest endorsement by the largest percentages of providers were that the COS process adds to much additional paperwork (32%) and that is takes time away from other important activities (31%).

One limitation of the data from Study 1 is that the data are based on providers' self-reported information about their assessments of their own and their other team members knowledge of and competencies in the COS ratings process. As such, the data may overestimate providers actual knowledge and competencies with the COS rating process. Additionally such self-reported data may also overestimate the accuracy of the ratings and the quality and integrity of the COS rating process because providers may have been reluctant to give responses to survey items that would reflect poorly on their own practices.

Section 4 —Study 2: Comparison with Child Assessments Study – Design, Methods, Key Findings

Research Questions for Study 2

The focus of the comparison with child assessments study was to examine the relationships between the COS and existing assessment tools. Specifically, the COS was compared to the Battelle Developmental Inventory, Second Edition (BDI-2; Newborg, 2005)and the second edition of the Vineland Scales of Adaptive Behavior (Vineland-II; Sparrow, Cicchetti, & Balla, 2005). Study 2 addressed three major research questions:

1. What are the descriptive characteristics of entry and exit COS ratings and assessment scores?
2. What is the relationship between COS ratings and scores on the BDI-2 and Vineland-II, at program entry and program exit?
3. To what extent do COS ratings and two multi-domain child assessment measures assign children's progress to the same OSEP reporting category?

Descriptive information on children's functioning also was collected with The ABILITIES Index (Bailey et al., 1995). This measure provides some additional information about the global impact of the child's disability on functioning across an array of different areas. Study 2 also examined the relationships between the ABILITIES Index and scores on the two assessment tools and COS ratings.

Assumptions in Study Design for Study 2

The rationale for examining the relationships between the COS and these tools extends from a psychometric validation approach in which a new tool is compared to an existing tool that measures the same construct. However, using this approach for validating the COS is challenging because there is no existing assessment tool that measures the three child outcomes. No tool, including the two child assessment tools chosen for use in the child assessments study, has been designed and validated to capture children's functioning across settings and situations on these outcomes. Therefore, as we consider evidence from comparisons of the COS to existing assessment tools, it is important to weigh the findings in the context of reasons why we expect relationships and agreement and reasons why we might also expect important differences.

It seems reasonable to expect that the COS and the BDI-2 and Vineland-II might show agreement and moderately strong relationships with one another. This expectation is based on the following facts:

- scores are based on assessment of the same child at approximately the same time frame,
- the same family provides input about the child's functioning to both the assessor and to team members involved in identifying a rating, and
- although not exactly the same, there is certainly similarity between the broad content of the three outcomes and what is measured by the BDI-2 and the Vineland-II.

Equally important are the reasons to expect some amount of disagreement among scores on the various approaches. Some of these reasons include:

- The BDI-2 and Vineland-II do not measure the same constructs as the COS.
- The BDI-2 and the Vineland-II are organized around domains, whereas the COS is organized around outcomes that intentionally integrate skills and functioning levels across domains.
- There are differences between approaches in the emphasis on examining functional use of skills across multiple settings and situations.
- The COS rating was designed to be based on varied experiences and assessments from multiple sources of information over a relatively brief period of time, whereas the BDI-2 and the Vineland-II contain single sources of information.
- Although efforts were made to minimize it, there was necessarily a time lag of usually three to five weeks between the BDI-2 and Vineland-II assessments (usually conducted at the same time) and the COS ratings. Infants and young children show rapid growth and variability in performance (which lowers test-retest reliability) even when the same assessment tool is used and,
- Guidance for using the COS instructs teams to examine children's functioning in everyday environments using whatever assistive technology is ordinarily available to them in those environments, whereas other assessment tools look for performance on a set of specific skills that may or may not allow for assistive technology to be used.
- The concept of "typically developing" is critical to the OSEP measurement approach and is thus embedded in the COS rating. The score range for typical development on the BDI-2 or the Vineland-II is open to interpretation (e.g., above 1 standard deviation below the mean, 1.5 standard deviations, etc.).

In addition to examining the relationship between the COS rating and the assessment scores, we also present some data on the relationship between the BDI-2 and the Vineland-II. The BDI-2 and Vineland-II were conducted by the same assessor during the same visit and are both organized around domains, so one would expect similarities between these approaches to be stronger than similarities between the COS and either of the assessment tools.

Sample for Study 2

Sample of Children and Families

Exhibit 4.1 presents information about the final sample of 153 children in the child assessments study. For various reasons, we were not able to recruit the number of children that we had intended to recruit. Even though they agreed to be in the study, some sites did not recruit the 6 families requested. To compensate, we asked other sites to recruit more than six families, but this still was not enough to get to the original goal of 206 children.

Entry Sample (n = 153)

The child assessments study draws its sample from local programs participating in ENHANCE. Recruitment of programs was described earlier in the general methods section (Section 2). Across the programs, the project received names of 198 children whose families consented to participate in the study. Thirteen of these children did not meet eligibility criteria, leaving 185 children and their families for possible enrollment in the study. Eight families withdrew from the study either before or after entry data were collected; all of their data were excluded from analysis. Families who withdrew cited concerns about time for the extra assessments within the timeframe as the primary reason.

Usable enrollment data were collected from 153 of the remaining 177 children. These children were drawn from 35 programs in 8 states and are referred to throughout the report as the "Entry Sample" (*n* = 153) (Exhibit 4.1). Children were enrolled in the child assessments study between September 2010 and November 2013.

Reasons varied for the lack of data on the 24 children whom the project attempted to enroll but were not included in the final entry sample. In most cases, assessors simply could not complete the assessment in the required time frame.[8] Reasons for delays included child hospitalization and major parent illness; however, in most cases the primary reasons that assessments were not completed were non-response from families, refusals to schedule assessments prior to needed dates, or delays from repeated rescheduling of assessments based on "no shows," family needs, or child illness.

Longitudinal Sample (n = 70)

The child assessments study followed children from their entry in EI or ECSE until the time when they exited the program, up to around 3 years after entry. Exits were expected to usually occur near age 3 for EI programs and around kindergarten entry for most ECSE programs. Project staff stayed in close contact with programs, with monthly phone calls or email check-ins

---

[8]   Assessments needed to be conducted in close proximity to the date teams decided on COS ratings. The number of days that were available for assessors to complete the assessment was sometimes rather short, especially if there were delays between the COS meeting date and when the program forwarded the family's consent information to SRI.

to follow the participating children and ensure information was shared about when the children were exiting the program.

Out of the original 153 children enrolled, 70 children had usable exit data. At the time of exit assessments:

- 21 children became ineligible for the study between entry and exit.
  - 11 of these children moved, left the program, went into foster care, or the program lost all contact with the family, so there was not comparable entry and exit COS data from the same program.
  - 1 child was terminally ill and the program did not conduct the usual exit procedures including the COS with the family.
  - 9 children left the program for a variety of reasons before 6 months passed, so the program was not required to complete a COS form.
- 18 additional children did not exit the program before the study terminated, so exit data were not available.

For 24 of the remaining 114 children for whom exit assessment data would be expected, ENHANCE did not receive information about the child exiting in time to conduct assessments within the required time frame. In most cases, these children exited early and program data were not updated early enough by providers to allow program staff to share timely exit information with the project.

Of the 90 children for whom exit information was received within the specified window, 15 assessments were not completed despite repeated attempts to schedule with the family within the assessment time frame.

Of the 75 assessments completed, data from five children could not be used.
- 2 had final dates between the COS and the assessment data that were too far apart.
- 3 had no exit COS available from the program despite numerous attempts to obtain it.

In summary, the final sample with complete COS, BDI-2, Vineland-II, and ABILITIES Index data for analysis is 70 children; throughout this document we refer to this group as the "Longitudinal Sample" (*n* = 70) (Exhibit 4.1). These children were drawn from 25 programs in 8 states. Additional information about factors influencing data collection was presented earlier under challenges and contextual features in Section 2 of this report.

Characteristics of Children in Samples

The children in the child assessments study sample were similar to many samples of young children with disabilities, and the longitudinal sample had few statistically significant differences from the entry sample of the study. Some key things to know about the sample include the following:

- Consistent with the population of young children enrolled in EI and ECSE, children in the study were more likely to be male than female ($p < .001$).
- The mean age at entry into the study was 25.5 months for the entry sample and 27.0 months for the longitudinal sample; this difference did not reach statistical significance.
- There were no statistically significant differences in age at entry among children in EI; 56% entered at 1 or 2 years of age in the entry sample and 63% entered at these ages in the longitudinal sample. The remainder entered at less than 1 year of age.
- Overall in EI, between two-thirds and three-fourths of the sample entered with a developmental delay and the remainder entered with a diagnosed condition.
- In ECSE, about two-thirds of the entry sample entered the study near or at 3 years of age, compared with about half of the longitudinal sample.
- The disability for about two-fifths of the entry sample and one-third of the longitudinal sample was a developmental delay. Speech or language impairment was the type of disability for 41% of the entry sample and 46% of the longitudinal sample. Approximately 20% in both samples had some other disability, with most diagnosed with autism or a disorder on the spectrum. The remainder of children with other types of disabilities had health impairments.

Exhibit 4.1   Characteristics of the Child Sample for Child Assessments Study

Characteristics	Entry Sample			Longitudinal Sample		
	EI (*n* = 95) n (%)	ECSE (*n* = 58) n (%)	Overall (*n* = 153) n (%)	EI (*n* = 46) n (%)	ECSE (*n* =24) n (%)	Overall (*n* = 70) n (%)
Geographic Regions						
Number of States Represented	7	6	8	7	5	8
Number of Programs Represented	19	16	35	16	9	25
Gender						
Female	42 (44%)	17 (29%)	59 (39%)	24 (52%)	8 (33%)	32 (46%)
Male	53 (56%)	41 (71%)	94 (61%)	22 (48%)	16 (67%)	38 (54%)
Race/Ethnicity						
African-American/Black	17 (18%)	9 (16%)	26 (17%)	7 (15%)	3 (13%)	10 (14%)
Hispanic/Latino	13 (14%)	8 (14%)	21 (14%)	9 (20%)	4 (17%)	13 (19%)
Caucasian	61 (64%)	37 (64%)	98 (64%)	28 (61%)	15 (63%)	43 (61%)
Other	4 (4%)	4 (7%)	8 (5%)	2 (4%)	2 (8%)	4 (6%)
Age at entry						
Mean (SD) in months	14.5 (9.8)	43.6**** (7.8)	25.5 (16.8)	16.3 (10.1)	47.3**** (7.7)	27.0 (17.5)
< 6 months	28 (29%)	0 (0%)	28 (18%)	10 (22%)	0 (0%)	10 (14%)
6-11 months	14 (15%)	0 (0%)	14 (9%)	7 (15%)	0 (0%)	7 (10%)
1 year	27 (28%)	0 (0%)	27 (18%)	11 (24%)	0 (0%)	11 (16%)
2 years	26 (27%)	4 (7%)	30 (20%)	18 (39%)	0 (0%)	18 (26%)
3 years	0 (0%)	37 (64%)	37 (24%)	0 (0%)	12 (50%)	12 (17%)
4 years	0 (0%)	15 (26%)	15 (10%)	0 (0%)	11 (46%)	11 (16%)
5 years	0 (0%)	2 (3%)	2 (1%)	0 (0%)	1 (4%)	1 (1%)
Disability type						
Diagnosed condition	29 (31%)	0 (0%)	29 (19%)	12 (26%)	0 (0%)	12 (17%)
Developmental delay	66 (69%)	23 (40%)	89 (58%)	34 (74%)	8 (33%)	42 (60%)
Speech or Language Impairment	0 (%)	24 (41%)	24 (16%)	0 (0%)	11 (46%)	11 (16%)
Other	0 (%)	11 (19%)	11 (7%)	0 (0%)	5 (21%)	5 (7%)

Exhibit 4.1   Characteristics of the Child Sample for Child Assessments Study (concluded)

	Entry Sample			Longitudinal Sample		
Characteristics	**EI** (*n* = 95) n (%)	**ECSE** (*n* = 58) n (%)	**Overall** (*n* = 153) n (%)	**EI** (*n* = 46) n (%)	**ECSE** (*n* = 24) n (%)	**Overall** (*n* = 70) n (%)
Level of Functioning – Reversed Severity of Disability/ABILITIES Index Total Score						
Mean (SD)	102.0 (11.4)	99.9 (10.1)	101.2 (11.0)	103.7 (8.8)	100.5 (9.4)	102.6 (9.1)
Low (< 100)	25 (26%)	25 (43%)	50 (33%)	9 (20%)	11 (46%)	20 (29%)
Moderate (100-107)	33 (35%)	17 (29%)	50 (33%)	18 (39%)	6 (25%)	24 (34%)
High (108-114)	37 (39%)	16 (28%)	53 (35%)	19 (41%)	7 (29%)	26 (37%)
Length of Time in Program (Between COS Entry and COS Exit)						
Mean (SD) in months	N/A	N/A	N/A	14.8 (9.0)	21.4** (10.5)	16.1 (9.4)
< 9 months	N/A	N/A	N/A	9 (20%)	7 (29%)	16 (23%)
9-11 months	N/A	N/A	N/A	19 (41%)	0 (0%)	19 (27%)
1 to 2 years	N/A	N/A	N/A	11 (24%)	7 (29%)	18 (26%)
2 years or more	N/A	N/A	N/A	7 (15%)	10 (42%)	17 (24%)

* *p* < .05, ** *p* < .01, *** *p* < .001, **** *p* < .0001.

Note: Six programs in the Entry Sample and two programs in the Longitudinal Sample are counted twice in the overall numbers because they provided data for both EI and ECSE children.

Note: Length of time in program is a rough approximation. Time shown is the time between COS entry rating and COS exit rating. However, programs varied in how soon after entry or how close to transitioning out of the program the COS was completed. Children actually received services somewhat longer than this time period suggests, but the specific length of service is not known.

Information was gathered using the ABILITIES Index to better understand the abilities and functioning of children in the study sample. Based on scores from the ABILIITES Index, a range of functioning was evident in the sample, although relatively few children had profound impairments across multiple areas. Challenges in functioning were often limited to a few areas and were more moderate in nature. In both the entry and longitudinal samples, approximately one-third of the children were in each of the three ABILITIES Index groupings.

In the longitudinal sample, an average of 16.1 months elapsed between the entry and exit COS ratings. The length of time children were in the program was fairly evenly distributed.

Sample of COS Teams

No specific criteria were specified about characteristics of COS teams to be included in the sample. Programs were instructed to complete the COS according to their usual protocols and with the team members who are normally involved for the children whose families consented to participate. Exhibit 4.2 shows information about the teams who completed entry and exit COS forms in both the entry and longitudinal samples.

Exhibit 4.2   Descriptive Information about Teams Completing the COS in Entry and Longitudinal Samples

Characteristics	Entry Sample			Longitudinal Sample		
	EI (*n* = 95) *n* (%)	ECSE (*n* = 58) *n* (%)	Overall (*n* = 153) *n* (%)	EI (*n* = 46) *n* (%)	ECSE (*n* =24) *n* (%)	Overall (*n* = 70) *n* (%)
Total Number on COS Decision Team (Professionals and Family Members)						
At Entry Mean (SD) in months	3.6 (1.5)	4.5 (2.2)	3.9 (1.9)	3.5 (1.6)	4.2 (2.4)	3.8 (1.9)
At Exit Mean (SD) in months	N/A	N/A	N/A	2.7 (1.4)	3.3 (1.9)	2.9 (1.6)
Family Member Present for COS Decision						
Yes at Entry	47 (51%)	30 (57%)	77 (53%)	22 (48%)	14 (58%)	36 (51%)
Yes at Exit	N/A	N/A	N/A	20 (43%)	10 (42%)	30 (43%)
Family Member Involvement Across Entry-Exit						
Family member not present at either COS	N/A	N/A	N/A	20 (43%)	9 (38%)	29 (41%)
Family member at entry COS only	N/A	N/A	N/A	6 (13%)	5 (21%)	11 (16%)
Family member at exit COS only	N/A	N/A	N/A	4 (9%)	1 (4%)	5 (7%)
Family member at both entry and exit COS	N/A	N/A	N/A	16 (35%)	9 (38%)	25 (36%)
Unknown	N/A	N/A	N/A	N/A	N/A	N/A
Service Coordinator on Team						
Yes at Entry	87 (92%)	43 (74%)	130 (85%)	42 (91%)	17* (71%)	59 (84%)
Yes at Exit	N/A	N/A	N/A	35 (81%)	13* (59%)	48 (74%)

* $p < .05$, ** $p < .01$, *** $p < .001$, **** $p < .0001$.

Note: Eight forms at entry and five forms at exit had incomplete information about the precise number and/or roles of COS team members. Percentages are based on the remainder in the sample.

Methods for Study 2

Measures

Data collected in the child assessments study included the second editions of the Battelle Developmental Inventory (BDI-2), the Vineland Adaptive Behavior Scales (Vineland-II) Survey Interview Form, the ABILITIES Index ratings, and the Child Outcomes Summary form (COS).

The BDI-2 (Newborg, 2005) is an assessment tool that is widely administered to young children with disabilities, especially as part of comprehensive assessments around eligibility or for research on these populations. The BDI-2 assesses a child's functioning through a combination of structured items administered to the child, unstructured items based on observation of the child in a natural setting, and caregiver report about skills that are not otherwise easily observed. The tool provides scores in five subdomains: Personal-Social, Communication, Cognitive, Motor, and Adaptive. The tool relies on the use of basal and ceiling rules associated with children of various ages for scoring. Psychometric information for the BDI-2 is well established and is provided in Appendix F. Standard scores (M = 100, SD = 15) based on a broad sample of children in the population, were used for the analysis. The BDI-2 was not normed specifically on a sample of children with disabilities.

The Vineland-II relies heavily on a face-to-face interview with the caregiver, supplementing that with some observation and direct assessment on specific items (Sparrow et al., 2005). The Vineland-II is a widely used tool, designed to provide greater detail about children's adaptive behavior than many other assessment tools. The Vineland II provides scores for Socialization, Communication, Motor, and Daily Living Skills domains. Psychometric information for the Vineland-II is well established and is provided Appendix F. The tool relies on the use of basal and ceiling rules associated with children of various ages for scoring. Standard scores (M = 100, SD = 15) were used for the analysis.

The ABILITIES Index is a profile of the child's abilities across 9 different areas(Bailey et al., 1995; Buysse, Smith, Bailey, & Simeonsson, 1993). The ABILITIES Index was designed to have someone who is very familiar with the child's functioning complete the measure, rating the child's functioning on a scale from 1 to 6 in each of the 9 areas (hearing/audition, behavior and social skills, intellectual functioning, intentional communication, limbs (use of hands, arms, and legs), tonicity, integrity of physical health, eyes/vision, and structural status). Ratings of 1 represent normal functioning for the child's age, or good health. Ratings of 6 indicate profound disability or extreme interference in functioning or health problems. Several areas are subdivided for separate ratings, resulting in 19 ratings across the 9 areas of functioning. For example right and left ear hearing is rated separately and left and right hands, arms, and legs are rated separately. Scores range from 19 – 114 with higher scores indicating greater interference with functioning. However, for analyses in the ENHANCE project, we have reversed scores so that

higher values represent higher functioning and lower values represent greater difficulties with functioning.

The COS process was described earlier and the forms are provided in Appendix A. The COS includes a rating from 1-7 and a single question about if the child made any progress between entry and exit for each of the three child outcomes areas. These features of the form are constant across programs, however, each local program used the version as implemented by the state in which they are located. In some cases, forms were formatted differently, inserted into other documents (e.g., the IFSP/IEP or assessment documents), and varied in the amount of space and requested information for documenting evidence that went with the form. More information about how various subdomains on assessments are expected to relate to COS ratings on the three outcomes areas is provided below in the data analysis section (Mapping of Domains to Outcomes).

Descriptive information about the child was gathered from the program using the Child and Family Information Form. The form included information about the child's disability as well as basic information for contacting the family and identifying the child's age and entry to and expected exit from the program. A copy of this form is included in Appendix H.

Methods for Training and Ensuring Fidelity of Assessment Tools

SRI contracted with skilled professionals in each of the geographic areas where programs were located to administer the BDI-2 and the Vineland-II. Individuals were identified who had experience working with young children with disabilities, conducting developmental assessments, conducting interviews and home visits with families, and who had considerable background knowledge in child development. Assessors had varied backgrounds in areas including early childhood, early childhood special education, social work, nursing, speech pathology, and educational psychology. All 17 assessors initially identified were flown to a common location and participated in a 3-day face-to-face training on the BDI-2 and Vineland-II. Topics covered included general background about the project, the assessor's role, paperwork and steps in completing an assessment visit, best practices for home-based assessments and interviewing families, calculating child age, basal and ceiling rules, scoring assessments, detailed coverage of the specific assessment tools and specific items with criteria, things to look for, and mistakes to avoid. Individuals were given opportunities to practice, especially with challenging interview and direct administration items, and feedback was shared about specific wording and criteria, administration errors, as well as tips and suggestions.[9]

After the initial training, assessors conducted practice assessments with non-study families who had children in the target age range of the program from which the children to be assessed

---

[9] Over time, 11 additional assessors were needed to replace original assessors who left the project or were not certified to do assessments, or as new sites were recruited. These assessors received a slightly modified version of the training described above; it was usually conducted via a series of webinars.

would be recruited. Practice administration the assessment protocols included children of varying ages, usually birth through age five. If the assessor was in an area where only EI or only ECSE programs were participating in the project, then the assessor practiced with children and families of the relevant age range. Assessors video-recorded the assessment sessions and submitted the video recordings and assessment protocols to project staff for review. Each consecutive video session underwent a thorough review and coding. Study staff provided feedback on videos and paperwork submitted, including administration technique, scoring, and general rapport with families, as well as accuracy in calculating age, scoring, use of basal and ceiling rules, and completeness of assessment tool forms Assessors repeated practice sessions and videos until they met the criteria for conducting research assessments. The assessors were required to demonstrate across a minimum of 2 complete sessions that they were competent to administer the assessments in a standardized format. Once an assessor met the criteria necessary for research assessments, she/he was deemed eligible to assess children in the study.[10] Assessors received ongoing communication with feedback about issues noted across assessments and a list of responses to frequently asked questions.

When research assessments were submitted, study staff reviewed assessment scoring protocols for completeness, appropriate computation of the child's age, following basal/ceiling rules, and scoring. Two study staff members reviewed each assessment and double-entered the data. Any concerns with the documentation, administration, or scoring were discussed by both study staff and resolved, contacting the assessor as needed.

Data Collection Methods

Programs in the study sample recruited families for participation in the child assessments study. Programs were encouraged to recruit families representing an array of ages at entry and with disability types in each of the categories identified for EI and ECSE. To participate, the children and families were required to meet the following criteria:

- Child is eligible to receive EI or ECSE services,
- Child is likely to receive services for at least 6 months and to stay in the area until their natural exit from the EI or ECSE program,
- Child has a parent/guardian who speaks English; if the child understands spoken language, they must understand English well enough for an assessment in English to be a valid reflection of their functioning,
- Child is not in foster care, and
- Child does not have a sibling who is already participating in the study.

Programs recruited families and collected informed consent. Programs sent the consent, family and child information form, ABILITIES Index, and COS process form to the SRI research team.

---

[10] Despite repeated efforts to support assessors in conducting standardized administrations, over the life of the project eight assessors were trained who did not meet study criteria to conduct the assessments. Training and support of assessors required many more resources than projected in the proposal.

Local programs identified an individual provider who knew the child well to complete the ABILITIES Index profile near the child's program entry. Usually the child's primary service provider or the lead professional on the child's eligibility assessment team completed the ABILITIES Index shortly after the child's family enrolled in the study. In most cases there was some delay between when program staff completed the COS and the ABILITIES Index. On average the ABILITIES Index was completed 14.0 (SD = 24.8) days after the COS in the entry sample and 15.2 (SD = 24.1) days after the COS at exit in the longitudinal sample.

Once SRI became aware that the family signed informed consent for participation, a study assessor was assigned to contact the family and administer the BDI-2 and Vineland-II to the child. The assessor was masked (i.e., the study ensured the assessor remained uninformed about) from information about the child's COS rating, ABILITIES Index ratings, and any of the team discussion related to the child's functioning. The assessor received the name and contact information for the child/family, the child's birth date, the name of the EI or ECSE program the child was enrolled through, and the child's disability. Assessment administration was conducted in the child's home, usually collecting data for both the BDI-2 and Vineland-II in the same day, at a time convenient to the family when the child would be able to demonstrate his/her skills. Assessors checked completed assessments for accuracy and sent them to SRI for final scoring and data entry.

Assessments were completed an average of 25 days after the COS process at entry (Entry Sample M = 25.4, SD = 11.1, range 0-59; Longitudinal Sample M = 24.9, SD = 10.3, range 1-41) and an average of 25.8 days (SD = 12.7, range 3-66) after the COS process at exit in the longitudinal sample. No significant differences were observed in the length of time between completion of the COS and the standardized assessment between EI and ECSE in either sample or between the entry and longitudinal samples. Families received a children's book and a $10 gift certificate at each assessment. Each family also was entered in an annual drawing, with one family per year receiving a $100 gift certificate.

SRI staff worked closely with the programs to be informed when children participating in the child assessments study were leaving the program and to schedule assessments with the family shortly after the expected COS meeting date. Because children could stay in the program 6 months to 3 years, and children were enrolled from September 2010 through November 2013, data collection extended over a number of years. Children in the final longitudinal sample exited between May 2011 and March 2014.

Data Analysis for Study 2

Mapping of Domains to Outcomes

The BDI-2 and Vineland-II both produce subdomain scores whereas the COS process does not produce information at the domain level. Rather, COS ratings are based on functioning on the three child outcomes: children have positive social relationships, children acquire and use

knowledge and skills, and children take appropriate action to meet their needs. As noted in Section 1, the outcomes focus on children's functioning to accomplish tasks that are meaningful to the child, and accomplishing daily tasks requires integrating skills across domains(Early Childhood Outcomes Center, 2005). So, for instance, communication or language is a domain on many assessment tools. However, communication is important to having positive social relationships, acquiring knowledge and skills, and being able to get one's needs met.

The theoretical structure of the COS implies that scores on a domain-based assessment may inform performance on multiple outcomes rather than having a one-to-one relationship with only one outcome. Because of this, we would not expect to see the usual pattern sought in validation studies, where one outcome would map completely with one target domain with little or no relationship between that outcome and other domains or between the target domain and the other outcomes. This feature poses an analytic challenge because often assessment tools are validated by examining correlations or using factor analysis to show that the three child outcomes can be rotated and show unique variance from one another with regard to specific constructs or domains.

Even though there are problems associated with mapping specific domains to specific outcomes, the ECO Center has provided some information about how data gathered from assessment tools might inform COS ratings, since teams identifying COS ratings have domain-based assessment information about each child to consider as one type of information available for their rating decision (Early Childhood Technical Assistance Center, 2014; The Early Childhood Outcomes Center, 2006). Specifically, for standardized assessment tools where it is not appropriate to examine item level scores or clusters smaller than at the subdomain level, guidance was provided about which outcome area most of the content on the assessment provides information about. Using this logic and guidance, data analysis focused on the following expected relationships between the COS outcomes and the BDI-2 subdomains:

- Personal-Social subdomain on the BDI-2 will provide information for positive social relationships on the COS,
- Both the Communication and Cognitive subdomains on the BDI-2 will provide information for acquiring and using knowledge and skills on the COS, and
- The Adaptive subdomain on the BDI-2 will provide information for taking appropriate action to meet needs on the COS.

Likewise, data analysis focused on the following expected relationships between the COS outcomes and the Vineland-II subdomains:

- Socialization subdomain on the Vineland-II will provide information for positive social relationships on the COS,
- Communication subdomain on the Vineland-II will provide information for acquiring and using knowledge and skills on the COS, and
- The Daily Living Skills subdomain on the Vineland-II will provide information for taking appropriate action to meet needs on the COS.

One other domain assessed by both the BDI-2 and the Vineland-II and sometimes considered in relationship to the child outcomes is the Motor domain. Taking appropriate action to meet needs includes getting from place to place as well as getting needs met, using tools, and self-help activities. Motor skills support a child's ability to accomplish these tasks. However, the way motor skills are assessed on most tools does not involve a universal design approach, so that children who use a walker or a wheelchair, for instance, would not receive credit for mobility on the assessment tool. This distinction does not matter for the majority of children in EI or ECSE because they do not have motor impairments. However, for a subset who do, the motor score may or may not be a good indicator of how well the child is able to get needs met. In some portions of this chapter, including in reports of descriptive data and correlations, we provide the information about the Motor subdomain and/or relationships between the Motor subdomain and taking appropriate action to meet needs COS ratings.[11] However, on most analyses for Study 2, we have left the motor subdomain out of the computations because of this key conceptual distinction.

The data analysis also included consideration of scores on the ABILITIES Index:

- Throughout most analyses, we used a total score for the ABILITIES Index to provide an overall sense of the child's functioning and the degree to which the child's disability is influencing functioning. Internal consistency for the Entry Sample was run, indicating an alpha of .88.
- For some mean comparisons by outcome, we combined scores on specific content from the index to more closely represent the content addressed in the outcome area.
  - For positive social relationships, we combined ABILITIES Index scores on behavior and social skills and intentional communication (alpha = .86). The resulting score is referred to as social/communication.
  - For acquiring and using knowledge and skills, we combined intellectual functioning and intentional communication (alpha = .85). The resulting score is referred to as cognitive/communication.
  - For taking appropriate action to meet needs, we combined ABILITIES Index scores on the limbs items, structural status, and tonicity (alpha = .92). The resulting score is referred to as structural integrity.
  - As with the total score, scores for the combined scores have been reversed such that higher values represent a higher level of functioning.

---

[11] A subgroup of states who do not use COS ratings for reporting accountability data to OSEP use the BDI-2. Some of these states consider the motor variable in algorithms to translate scores to progress categories while others do not. As a result, where possible, we have provided data both ways for reference to the field.

Assigning Progress Categories

The background section (Section 1) reviewed information about what child outcomes data state EI and ECSE programs report to OSEP. Statewide data are reported as numbers and percentages in each of five progress categories that sum to 100% of children in the state. A simple mathematical algorithm then converts the five progress categories into two summary statements capturing the percentage of children in the group who (1) demonstrate greater than expected growth, and (2) exit the program at age expectations.

The COS was designed to summarize the status of the child's functioning in each outcome and, by comparing entry and exit COS ratings, to classify each child's change in functioning into one of five OSEP progress categories. Given this important use of COS data for accountability purposes, analyses in Study 2 examine relationships between progress categories assigned based on the COS and progress categories based on the BDI-2 and Vineland-II. Some states use information from assessment tools directly to compute progress categories for reporting. The BDI-2 is one such tool used by some states, but neither the BDI-2 nor the Vineland-II was designed to assess functioning on the three child outcomes or for reporting OSEP progress categories. These assessment tools do not provide guidance about how to convert scores to progress categories, and states who use this approach use a variety of different algorithms. For comparative purposes, we classified data from these assessment tools into the progress categories using the criteria shown in Exhibit 4.3. Children had to perform within 1 standard deviation of mean functioning on all subdomains relative to the outcome to be considered functioning at age expectations at either entry or exit. Assessment scores needed to increase by 1/3 of a standard deviation (i.e., 5 points) or more on at least one subdomain related to the outcome in order to be classified as changing trajectory. In order to demonstrate any progress, children's assessment scores needed to show an increase in raw scores from one time to another even if their standard scores did not increase. The criteria for typical development and trajectory change are logical, but arbitrary. The use of different criteria for calculating progress categories would have led to different findings on the correspondence between COS and assessment score classifications on progress categories.

Exhibit 4.3   Decision Rules for Converting BDI-2 and Vineland-II Standardized Scores to OSEP Progress Categories

Progress Category	Label	Entry Standard Score	One or Both* Sub-domain(s)	Exit Standard Score	One or Both* Sub-domain(s)	Change Standard Score	One or Both* Sub-domain(s)	Raw Score Change	Comments
E**	**Functioning like same aged peers**	≥ 85	Both	≥ 85	Both				Enter and exit at age expectations is critical for this category.
D	**Improved functioning to that of same aged peers**	< 85	One	≥ 85	Both	≥ 5 pts	One		Show change in standard score, and, in at least one subdomain, begin below age expectations but cross over into the age-expected level(s).
C	**Moved closer to functioning like same aged peers**	No requirement, although almost all children in this group enter at < 85 on at least one subdomain		< 85	One	≥ 5 pts	One		Change in standard score >=5 on at least one subdomain is critical for this category. At exit, child must still function below age-expected levels on at least one subdomain.
B	**Improved functioning, no change in trajectory**	No requirement, as long as child does not enter and exit at age expectations (category e)		No requirement, as long as child does not enter and exit at age expectations (category e)		< 5 pts	Both	≥1 point or greater on both sub-domains	Raw score increase, but standard score changes by less than 5 points.
A	**Did not improve functioning**	No requirement, as long as child does not enter and exit at age expectations (category e)		< 85	One	< 5 pts	Both	< 1 point on both sub-domains	Not in category e and raw score stays the same or decreases.

* Where 2 or more subdomains from an instrument inform scores on the outcome rating, this column shows whether scores required to meet the progress category criteria, are based on achieving the threshold on only one of the subdomains or on both subdomains.

** Individuals who meet the pattern for progress category e are identified first and removed from consideration for progress categories a, b, c, and d.

Data Analyses

Data analyzed for Study 2 focused on standard descriptive analyses, including examining the numbers and percentages of children by category or mean values and standard deviations. Relationships between variables were examined by looking at correlations between each of the various assessment tools and the COS at a single point in time, and by undertaking ANOVAs between groups or linear regressions using the whole group to better understand relationships between variables. Statistical significance was tested using chi-squares, *t*-tests, and confidence intervals. As needed, additional detail about analytic methods is provided with the findings in sections below.

Key Results for Study 2: Comparison with Child Assessment Tools

Distributions of COS Ratings

In this section we present information about the distributions of the COS ratings for Study 2. Mean entry and exit COS ratings are provided, including means for children with varied levels of impairment. The next set of exhibits shows the percentages of children at each of the COS rating categories for each outcome. Then, exhibits examine the distribution of change from entry to exit in the longitudinal sample. Finally, we consider variations in COS ratings based on the child's characteristics.

Mean COS ratings

As shown in Exhibit 4.4, overall mean ratings of the COS at entry range from 4.1 to 4.6 across the three child outcomes. These values are not significantly different from entry ratings observed in the Longitudinal Sample ($n = 70$) (Exhibit 4.5).

Exhibit 4.4   Mean COS Ratings at Program Entry Overall and for EI and ECSE in Entry Sample ($n = 153$)

At Program Entry	Positive Social Relationships Mean (SD)	Knowledge and Skills Mean (SD)	Action to Meet Needs Mean (SD)
Overall ($n = 153$)	4.6 (1.6)	4.1 (1.5)	4.4 (1.5)
EI ($n = 95$)	4.8 (1.6)	4.2 (1.6)	4.3 (1.4)
ECSE ($n = 58$)	4.2 (1.6)	4.1 (1.4)	4.5 (1.6)

Exhibit 4.5   Mean COS Ratings in Longitudinal Sample Overall and for EI and ECSE at Entry and Exit

	Entry (*n* = 70)			Exit (*n* = 70)		
	Positive Social Relationships	**Knowledge and Skills**	**Action to Meet Needs**	**Positive Social Relationships**	**Knowledge and Skills**	**Action to Meet Needs**
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**
Overall (*n* = 70)	4.7 (1.6)	4.2 (1.6)	4.5 (1.6)	5.5 (1.5)	5.4 (1.4)	5.5 (1.6)
EI (*n* = 46)	4.9 (1.6)	4.1 (1.6)	4.4 (1.5)	5.7 (1.5)	5.4 (1.5)	5.4 (1.6)
ECSE (*n* = 24)	4.4 (1.8)	4.2 (1.6)	4.5 (1.8)	5.3 (1.5)	5.5 (1.3)	5.7 (1.5)

Percent of children at each COS rating category: Overall

Exhibit 4.6 shows the percent of children at each of the 1-7 rating categories on the COS at entry.

- At entry, there are children at each of the 7 rating points on all three outcomes.
- There are many fewer children with COS ratings in the 1-3 range than at the higher end of the scale.
- As predicted, less than 10% of children in the study received ratings of 1 on all three outcomes; this is consistent with the expected distribution of children's functioning at entry to EI and ECSE programs.

    Notably, more children entered at age-expected levels (i.e., COS ratings of 6 or 7) for positive social relationships than with the other outcomes.

Exhibit 4.6   Percent of Children at Each COS Rating Category across all Outcomes in Entry Sample
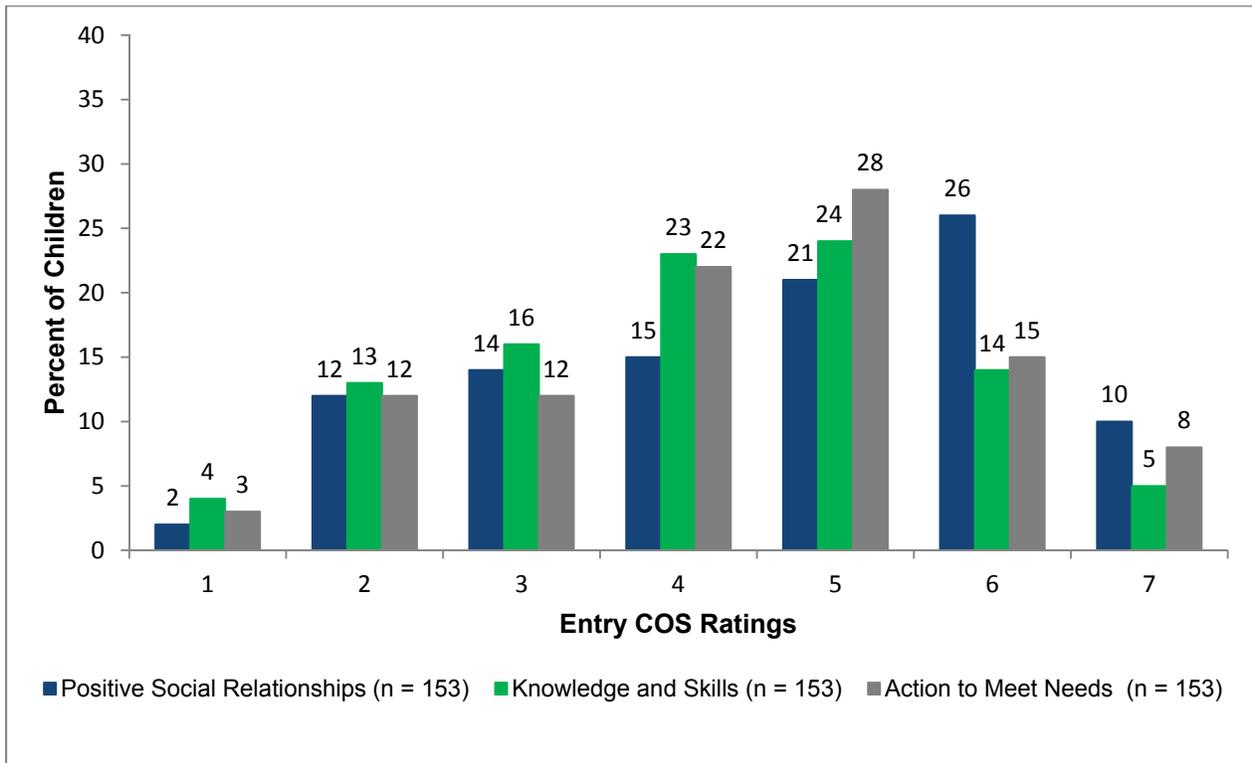
Percent of children at each COS rating category: At Entry and Exit

Exhibits 4.7, 4.8, and 4.9 show the percentage of children across COS rating categories for each of the three outcomes at both entry and exit in the longitudinal data ($n = 70$).

- The patterns are generally similar to those in the entry sample with a slightly higher percentage of children receiving 6's at entry on positive social relationships and knowledge and skills than in the entry sample.

- The longitudinal sample also shows all seven rating points used on the COS at entry and finds all seven rating points also in the distribution at exit.

- Like in the entry sample, the longitudinal sample data indicate that less than 10 percent of those in the sample received an entry rating of one on each of the outcomes.

- Exit distributions show a shift to the right with a higher percentage of children receiving age-expected ratings (i.e., 6's, and 7's) at exit compared to entry. This pattern is expected since the goal of EI and ECSE programs is to improve children's functioning.

- The shift in mean distributions from entry to exit in the longitudinal data from the child assessments study is consistent with higher mean levels at exit than at entry throughout extant data reported by both state EI and ECSE programs (Section 7, Study 4).

Exhibit 4.7   Percent of Children by entry and exit COS ratings on Positive Social Relationships in Longitudinal Sample (*n* = 70)
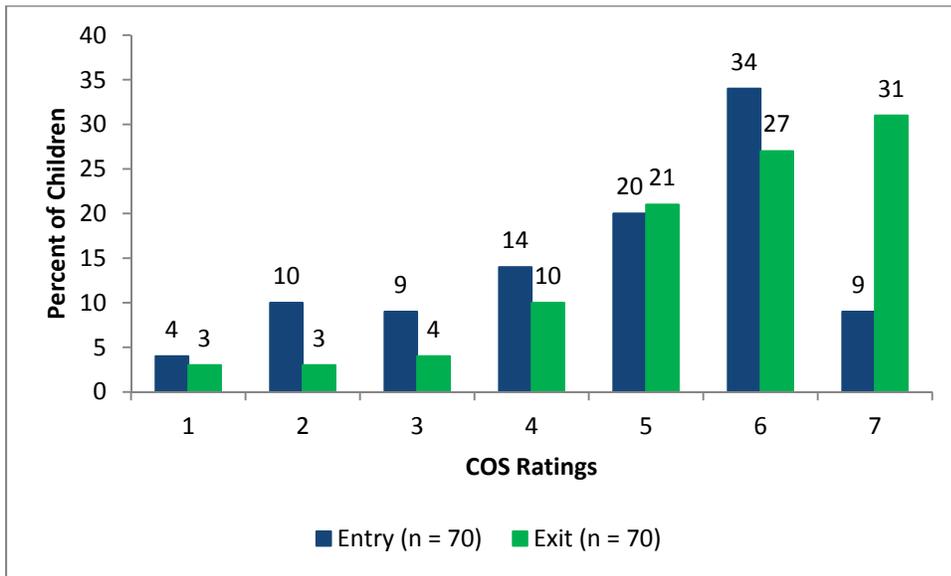
Exhibit 4.8   Percent of Children by entry and exit COS ratings on Acquiring and Using Knowledge and Skills in Longitudinal Sample (*n* = 70)
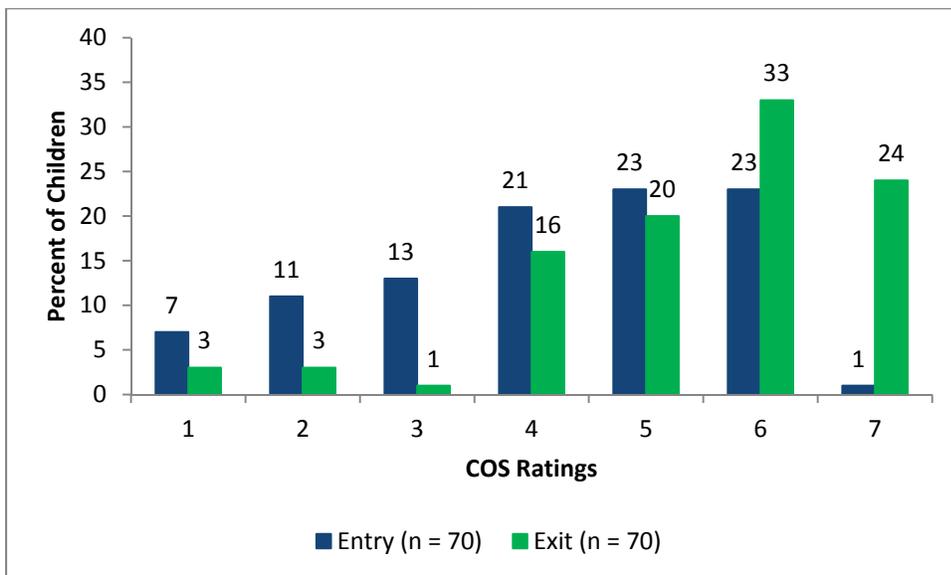
Exhibit 4.9   Percent of Children by Entry and Exit COS ratings on Taking Action to Meet Needs in Longitudinal Sample (*n* = 70)
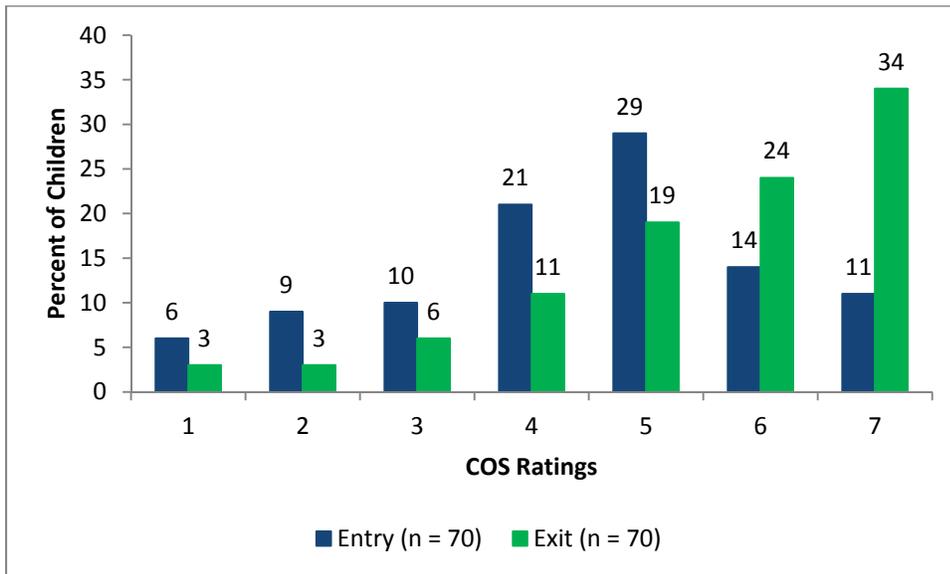
Percent of children at age-expected levels across all three outcomes

Exhibit 4.10 shows the percentage of children who enter EI or ECSE at age-expected levels (i.e., COS rating of 6 or 7) on all three outcomes. Information is provided for both the entry sample and the longitudinal sample.

- Ten percent of entry sample entered at age-expected levels on all three outcomes based on COS ratings. No national data are available on the percentage of children entering EI and ECSE who function at age-expected levels on all three outcomes.

- It also includes the percentage of children who would be at age-expected levels based on the BDI-2 scores or Vineland scores.

The ten percent of children identified by the COS as at age-expected levels on all three outcomes at entry also had higher average ABILITIES Index total scores (*M* = 108 [*SD* = 3.9] age-expected children versus *M* = 100 [*SD* = 11.2] for those not at age-expected levels on all three outcomes, *p* < .0001).

Exhibit 4.10 Children Entering at Age-Expected Levels on All Three Outcomes by Measurement Approach in both Entry and Longitudinal Samples

Percentage of Children Entering at Age Expectations on All Three Outcomes	Entry Sample (*n* = 153) n (%)	Longitudinal Sample (*n* = 70) n (%)
COS	16 (10%)	10 (14%)
BDI-2	23 (15%)	15 (21%)
Vineland-II	23 (15%)	15 (21%)
BDI-2 with motor	22 (14%)	14 (20%)
Vineland-II with motor	19 (12%)	12 (17%)

Distribution of change from entry to exit

Exhibits 4.11 – 4.13 display the extent of change in COS ratings between entry and exit in each of the three child outcomes.

- In each case, between one-quarter and one-third of longitudinal study sample received the same rating from entry to exit.
- Most children's ratings increased between entry to exit, with 40-51 percent of children's COS ratings increasing by either one or two rating points.
- Six to fifteen percent of children had COS ratings decrease between entry and exit.

As a reminder, the COS is based upon a scale that has distance from age-expected embedded within it. So, the expectations for the kinds of skills and behaviors used in order to be considered functioning at an age-expected level increases as the child's age increases.

- In order to maintain the same rating from entry to exit, the child has to display many new skills and "keep pace" with the trajectory of growth other children of the same age experience.
- Numerical increases represent shifts in trajectories, such that the child has acquired enough skills to narrow or close the gap between the child's functioning and what is expected at that age.
- Numerical decreases indicate that the child's rate of growth has slowed relative to the rate observed among same-aged peers, resulting in a greater distance between the child's functioning and what is expected at that age at exit than there was at entry[12]. Nearly all children, including most of those whose COS ratings decrease between entry and exit, still acquire and use new skills between entry and exit.

---

[12] Children with numerical decreases in COS ratings may still acquire new skills, however the rating decrease indicates that their set of skills at exit is farther from age-expected skills at exit than it was at entry. This means that the rate of individual growth observed among these children is slower than that the rate of growth observed during the same time period among their same-aged peers.

Exhibit 4.11   Change in COS Ratings from Entry to Exit in Positive Social Relationships (*n* = 70)
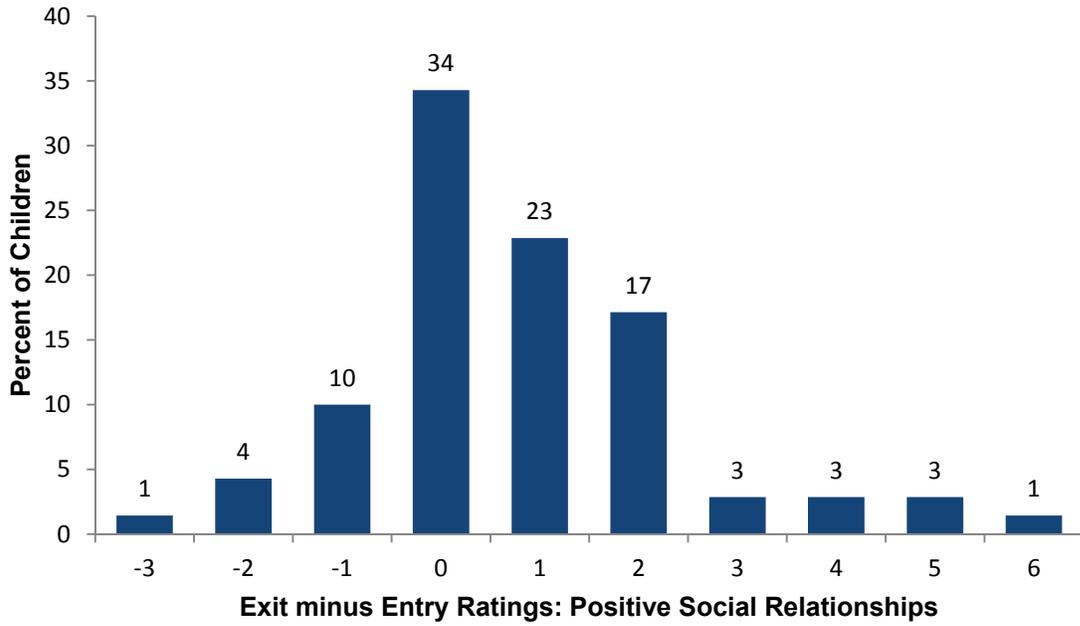
Exhibit 4.12   Change in COS Ratings from Entry to Exit in Acquiring and Using Knowledge and Skills (n = 70)
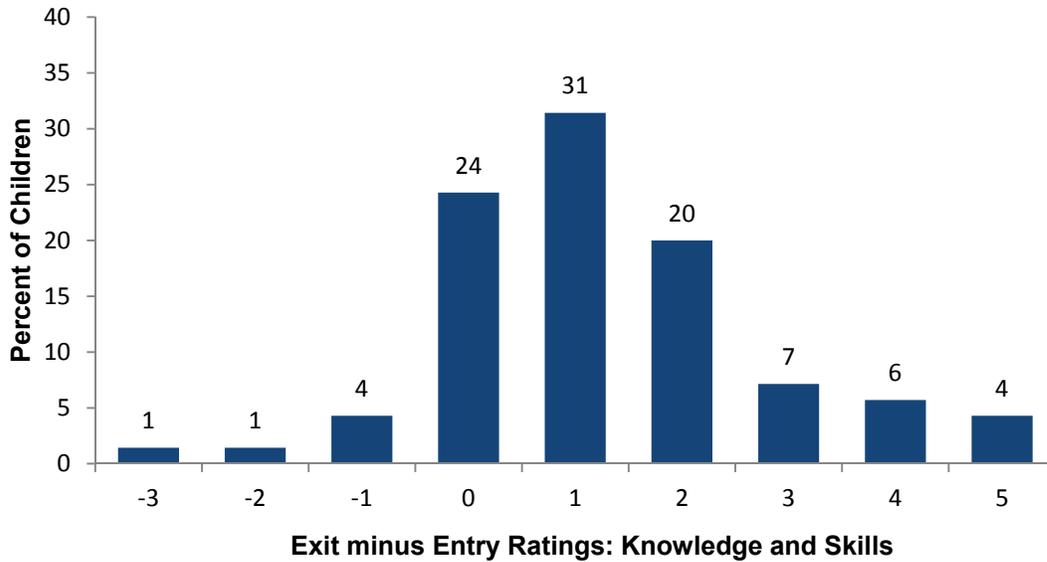
Exhibit 4.13   Change in COS Ratings from Entry to Exit in Action to Meet Needs (*n* = 70)
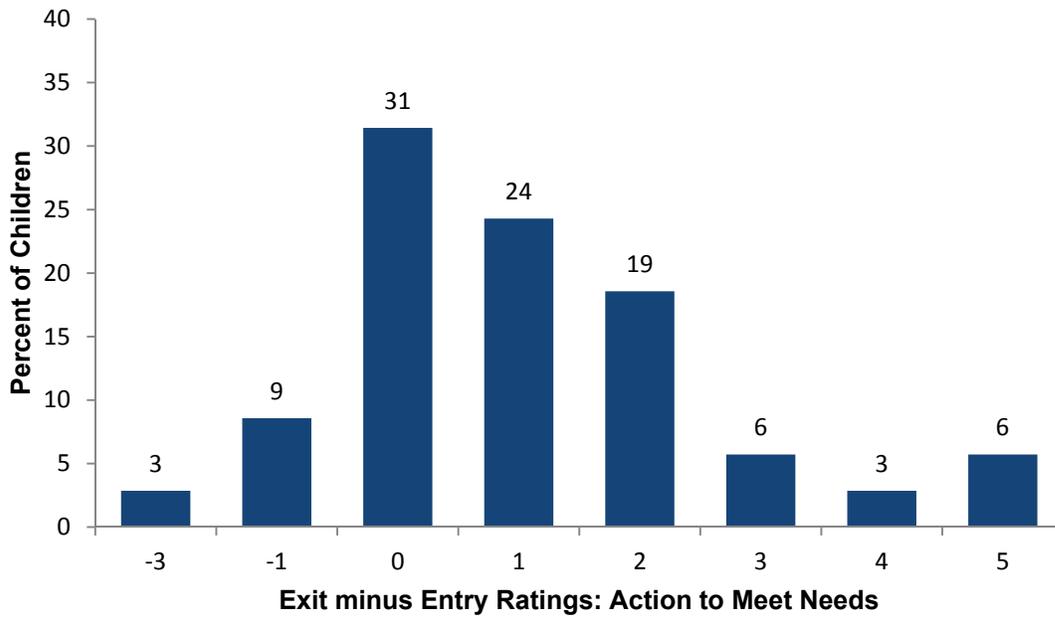
Given the amount of developmental growth required to show increases in COS ratings between entry and exit, an increase or decrease of 4 or more COS points is not expected to occur very often. Exhibit 4.14 shows the extent to which different size increases and decreases in functioning occurred in the longitudinal sample. Few children had 4 or more point increases for entry to exit for any of the outcomes (7%, 10%, 9%, respectively).

Exhibit 4.14   Summary of Extent of Change in COS Ratings Between Entry and Exit on All Three Outcomes in the Longitudinal Sample (*n* = 70)

Exit minus Entry Change Patterns (*n* = 70)	Positive Social Relationships *n* (%)	Knowledge and Skills *n* (%)	Action to Meet Needs *n* (%)
4 or more point decrease	0 (0%)	0 (0%)	0 (0%)
2-3 point decrease	4 (6%)	2 (3%)	2 (3%)
Stable or plus or minus 1 point	47 (67%)	42 (60%)	45 (64%)
2-3 point increase	14 (20%)	19 (27%)	17 (24%)
4 or more point increase	5 (7%)	7 (10%)	6 (9%)
Total	70 (100%)	70 (100%)	70 (100%)

Relationships were examined to see if there were clear patterns in characteristics of children with large COS rating changes. No clear pattern was observed in ratings originating from EI versus ECSE programs, particular programs, ABILITIES entry scores of the children, types of disability, child age at entry, length of time in the program, or whether or not a parent was involved in the team deciding on the COS rating. Review of BDI-2 and Vineland-II data found that children with four or more point increases on the COS also had higher scores on those tools over time, but not necessarily with as extreme changes in scores.

Variations in COS ratings based on the child's characteristics

Exhibit 4.15 provides information describing variation in COS ratings across children with different characteristics. Mean COS ratings are provided by disability type, by demographic characteristics, type of disability, and level of functioning. At entry into EI, there were no statistically significant differences in COS ratings between children with diagnosed conditions and children with developmental delays. Both of these categories contain children with a wide variety of conditions and severity levels so this is not surprising. At entry into ECSE, it was expected that children with speech or language impairments would have higher COS ratings at entry than children with either of the other disability types (Developmental Delay or Other). Statistically significant differences were found across all three outcome areas (all *p* <.001): children with speech-language impairments had higher average COS ratings at entry than did children with developmental delays or other types of disabilities.

Exhibit 4.15   Average COS Ratings by Program Type and Type of Disability in Entry Sample (*n* = 153)

Child's Disability Type at Program Entry	Positive Social Relationships		Knowledge and Skills		Taking Action to Meet Needs	
	Mean (SD)		Mean (SD)		Mean (SD)	
EI Programs						
Diagnosed Condition (*n* = 29)	4.6 (2.0)		4.2 (1.9)		4.2 (1.8)	
Developmental Delay (*n* = 66)	4.9 (1.4)		4.2 (1.4)		4.4 (1.3)	
ECSE Programs						
Speech or Language Impaired (*n* = 24)	5.0 (1.4)		4.8 (1.4)		5.5 (1.6)	
Developmental Delay (*n* = 23)	4.0 (1.3)	3.6*** (1.4)	3.8 (1.0)	3.5*** (1.2)	4.1 (1.3)	3.8*** (1.3)
Other condition (*n* = 11)	2.8 (1.3)		2.8 (1.3)		3.1 (1.1)	

* *p* < .05, ** *p* < .01, *** *p* < .001.

Exhibit 4.16 provides mean ratings on the COS at program entry for different groups of children by age and gender. No statistically significant differences were found in COS ratings on any of the outcomes based on the child's gender. The exhibit also provides additional information about average COS ratings for children of different ages in the sample.

Exhibit 4.16   Average COS Ratings Based by Gender and Age at Entry of the Entry Sample (*n* = 153)

	Entry Sample (*n* = 153)		
	Positive Social Relationships	Knowledge and Skills	Taking Action to Meet Needs
	Mean (SD)	Mean (SD)	Mean (SD)
Gender			
Male (*n* = 94)	4.6 (1.6)	4.2 (1.5)	4.5 (1.5)
Female (*n* = 59)	4.6 (1.7)	4.1 (1.6)	4.2 (1.5)
Age			
<1 year (*n* = 94)	4.9 (1.8)	4.5 (1.7)	4.1 (1.7)
1 year (*n* = 27)	4.8 (1.6)	4.1 (1.6)	4.1 (1.2)
2 years (*n* = 30)	4.6 (1.5)	3.9 (1.3)	4.7 (1.1)
3 years (*n* = 37)	4.1 (1..6)	3.9 (1.6)	4.4 (1.7)
4 years or older (*n* = 17)	4.6 (1.5)	4.4 (1.2)	4.8 (1.5)

Variations in COS ratings based on the child's level of functioning on the ABILITIES Index

COS ratings also were examined relative to the severity of the child's disability. Correlations between COS ratings and the ABILITIES Index scores were .54 for positive social relationships, .53 for knowledge and skills, and .59 for taking action to meet needs in the entry sample (n = 153) indicating moderate to strong relationships between the extent to which the child's the disability influenced the child's functioning as measured by the ABILITIES Index and the child's COS rating.

Mean COS ratings also were compared for the approximately 1/3 of children who scored below 100 on the ABILITIES Index scale versus for others with higher scores. In this case, lower scores indicate that the child's disability or delay was more severely impacting the child's functioning across a range of areas. Exhibit 4.16 shows these results. Statistically significant differences were found on the COS between groups such that children with higher levels of functioning also received significantly higher COS ratings on average across all three outcomes. Similar findings also were observed for each of the subdomains of the BDI-2 and Vineland-II across the outcomes. The only exception was BDI-2 communication. Although scores were higher on average for children with higher levels of functioning based on the ABILITIES Index, the difference was not statistically significant for the BDI-2 communication subdomain. Also, it is notable that while the expected patterns of variation generally occurred, the average standard score level for the group with higher levels of functioning based on the ABILITIES Index was not consistent across subdomains, fluctuating above the 1 standard deviation below the mean on the assessment tool on some subdomains and below that threshold on others.

Exhibit 4.16 Mean Entry COS Ratings and Mean Selected Domain Scores on the BDI-2 and Vineland-II for Children with ABILITIES Index Scores of Less than 100 Compared 100 or Higher from the Entry Sample (*n* = 153)

Mean Ratings or Standardized Scores (SD)	Positive Social Relationships		Knowledge and Skills		Action to Meet Needs	
	ABILITIES Index* < 100 (n = 50)	ABILITIES Index 100 or higher (n = 103)	ABILITIES Index* < 100 (n = 50)	ABILITIES Index 100 or higher (n = 103)	ABILITIES Index* < 100 (n = 50)	ABILITIES Index 100 or higher (n = 103)
COS	3.4 (1.51)	5.1** (1.35)	3.2 (1.57)	4.6** (1.28)	3.2 (1.42)	4.9** (1.22)
BDI-2 Personal-Social	79.1 (13.33)	85.7* (12.61)				
Vineland-II Socialization	73.6 (12.28)	79.4* (11.07)				
BDI-2 Communication			73.3 (16.37)	77.7 (15.86)		
BDI-2 Cognitive			70.9 (13.37)	81.3** (16.10)		
Vineland-II Communication			73.0 (15.57)	81.4* (12.03)		
BDI-2 Adaptive					76.3 (15.80)	85.4** (14.58)
Vineland-II Daily Living Skills					76.1 (13.02)	83.8** (11.53)
BDI-2 Motor					74.9 (13.87)	88.2** (15.32)
Vineland-II Motor					72.4 (11.40)	83.2** (11.01)

* $p < .01$, ** $p < .001$.

Note: ABILITIES Index scores have been reversed such that lower scores represent lower functioning and higher scores represent higher functioning.

Mean COS ratings at exit for those entering above and below age-expected levels at entry

Exhibit 4.17 shows the mean exit ratings for each outcome among those who entered at age expected levels (ratings of 6-7 on COS) and those who entered below age-expected ratings (ratings of 1-5 on the COS). On each outcome, children who entered at age-expected levels also exited at age expected levels.[13] The range of exit scores on the COS among children in these groups is also consistent with expectations given entry at or below age-expected levels.

---

[13] These children exited at age-expected levels. Both ratings of 6 and 7 on the COS represent currently functioning at age-expected levels. Ratings of 6 indicate that the team had some significant concerns about the child's functioning that might pose challenges for maintaining age-expectations over future. But current functioning remains at an age-appropriate level.

Exhibit 4.17   Mean Exit COS Ratings for Children Entering At versus Below Age Expected Levels on Each Outcome in Longitudinal Sample (*n* = 70)

Entry Rating (*n* = 70)	Exit Ratings on Same Outcome (*n* = 70)				
	N	Mean	SD	Minimum	Maximum
Positive Social Relationships					
COS 6-7 (at age expectations)	30	6.0	0.93	4	7
COS 1-5 (below age expectations)	40	5.1*	1.74	1	7
Knowledge and Skills					
COS 6-7 (at age expectations)	17	6.2	.90	4	7
COS 1-5 (below age expectations)	53	5.2***	1.49	1	7
Action to Meet Needs					
COS 6-7 (at age expectations)	18	6.4	.86	4	7
COS 1-5 (below age expectations)	52	5.2***	1.63	1	7

* *p* < .05, ** *p*< .01, *** *p* <.001.

Relationships between characteristics of child and COS Team with COS entry ratings

To understand how key characteristics of the child and COS team are related to COS Entry ratings, we conducted a series of regressions. In each case, we examined the extent to which each variable predicted entry COS ratings after controlling for the ABILITIES Index scores on the subareas related to the outcome. A summary of these results is provided below in Exhibit 4.18. Detailed results from these regressions are shown in Appendix J.

These regressions showed that after taking into account the child's level of functioning on the relevant ABILITIES Index subarea:

- Children in EI had lower COS ratings than those in ECSE for knowledge and skills (*p* < . 01).

- African-American children had lower COS ratings than white children for action to meet needs (p < . 05).

- Children who were older when they entered EI or ECSE programs had somewhat higher COS ratings than those who entered at younger ages for positive social relationships and knowledge and skills (both *p* < . 01).

Exhibit 4.18 Summary Findings From Regressions Predicting COS Ratings at Entry (*n* = 153)

Predictor Variable	Positive Social Relationships			Knowledge and Skills			Action to Meet Needs		
	B	SE B	β	B	SE B	β	B	SE B	β
Early Intervention Program	-0.32	.23	-0.10	-0.56**	.21	-0.18	0.04	.22	0.01
Male Gender	0.35	.21	0.11	0.31	.20	0.10	0.14	.22	0.05
Race/Ethnicity									
Hispanic	0.39	.30	0.08	0.17	.29	0.04	-0.34	.31	-0.08
African-American/ Black	-0.19	.28	-0.04	-0.47	.27	-0.12	-0.65*	.29	-0.16
Other Race/Ethnicity	0.19	.46	0.03	-0.19	.44	-0.03	-0.06	.48	-0.01
Age at Entry (Months)	0.02**	.01	0.21	0.02**	.01	0.20	0.00	.01	0.01
Parent in COS Rating Decision Meeting	0.14	.21	0.04	0.32	.19	0.10	0.34	.22	0.11
Number on COS Team	0.01	.06	0.01	-0.03	.05	-0.04	-0.01	.06	-0.01

* *p* < .05, ** *p* < .01, *** *p* < .001.

Note: Values shown are for predictor of COS entry ratings in model using outcome-relevant subareas of ABILITIES Index included as a covariate in every model.
In every model the ABILITIES Index subarea was a significant positive predictor of COS Entry ratings.

Relationships of COS Ratings with Assessment Tool Scores

An important research question focused on the relationships between the COS and the assessment tools. We considered relationships between the COS and the BDI-2 and Vineland-II with the largest available sample, the Entry Sample ($n = 153$) in this study.

Relationships at program entry: Levels of COS ratings with mean scores on the BDI-2, Vineland-II, and ABILITIES Index

The first issue addressed is the extent to which entry COS ratings for each outcome correspond to the assessment tool scores. The issue was examined by computing the mean scores on relevant subdomains of the BDI-2 and the Vineland-II for different COS rating points. Relationships also are shown for the ABILITIES Index subareas most relevant to each outcome. Although the ABILITIES Index is not a direct assessment tool, we expected ABILITIES scores to be related to COS ratings based on the information they provide about the level of functioning of children. The ABILITIES Index items were grouped into subareas so only those most likely to inform a specific outcome are presented. (See the earlier discussion under the data analysis in this section about the mapping of domains to outcomes and more information about the clustering of items used on the ABILITIES Index.)

We expected to see a stair-step pattern such that higher COS ratings would be associated with higher assessment tool scores. Results are shown in Exhibits 4.19 through 4.21 for each of the three outcomes. The sample size of 153 was too small to create reliable mean estimates at all 7 points on the scale, so the data described below show mean assessment tool scores at three levels of the COS: ratings on children of 1-3, 4-5, and 6-7. These groupings of ratings correspond to rating scale criteria in the following ways:

- children with foundational and immediate foundational skills only (ratings of 1-3),
- children with a mix of age-expected skills and skills that are below what is expected for the child's age (ratings of 4-5), and
- children functioning at age-expected levels on the outcome (ratings of 6-7).

Appendix K includes results for mean scores across all seven rating points. However, these data should be viewed with caution since mean estimates are based on very small sample sizes, especially for ratings of 1 and 7.

Overall, across the three assessment tools examined in relation to COS ratings on the outcomes, the expected stair-step pattern was found. Mean assessment tool scores for children with different COS ratings showed the expected patterns of increasing mean scores in 11 out of 12[14] of the mean scores calculated (92%) when COS ratings were clustered into 3 groups (1-3, 4-5, 6-7).

---

[14]   The stair-step pattern was found in the expected direction for 11 out of 12 comparisons. However, differences between each of the neighboring values did not all reach statistical significance. Lack of significance for some comparisons is likely due to the small sample sizes for some of the comparisons when the Entry Sample ($n = 153$) was split across 3 groups.

Positive Social Relationships

Across the assessment tools, for positive social relationships, the stair-step pattern held for the three levels of COS ratings (1-3, 4-5, and 6-7); higher mean scores were found for children with higher COS ratings (Exhibit 4.19). All of the differences between neighboring values reached a statistically significant threshold ($p < .05$) except for the difference between the first two bars on the Vineland-II Socialization subdomain ($p < .08$).

Exhibit 4.19   Mean Assessment Tool Scores by 3 Groups of COS Entry Ratings (1-3, 4-5, and 6-7) for Positive Social Relationships in Entry Sample ($n = 153$)
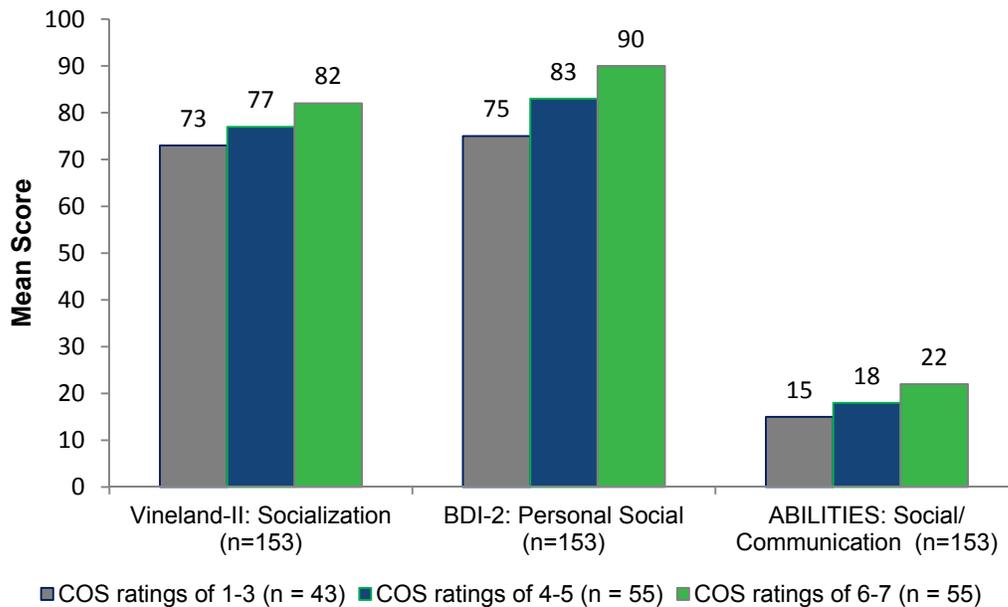
Knowledge and Skills

Across the assessment tools, for knowledge and skills, the stair-step pattern was found for most of the three levels of COS ratings (1-3, 4-5, and 6-7), such that children with higher COS ratings had higher mean scores on expected assessment subdomains (Exhibit 4.20).

- All of the differences between neighboring values reached a statistically significant threshold ($p < .05$) except for two of the comparisons involving the smallest group of children, the 30 children who received ratings of 6 or 7 at entry.
  - Differences on the Vineland-II Communication subdomain between children with COS ratings of 4-5 versus 6-7 did not visually show the stair-step pattern and did not reach statistical significance. However, the BDI-2 Communication subdomain did show a statistically significant difference in the expected direction.

o   Differences on the BDI-2 Cognitive subdomain between children with COS ratings of 4-5 and 6-7 showed the expected pattern but were not strong enough to reach the threshold for statistical significance.

Exhibit 4.20   Mean Assessment Tool Scores by 3 Groups of COS Entry Ratings (1-3, 4-5, and 6-7) for Knowledge and Skills in Entry Sample (*n* = 153)

Legend: ■ COS ratings of 1-3 (n = 51)   ■ COS ratings of 4-5 (n = 72)   ■ COS ratings of 6-7 (n = 30)
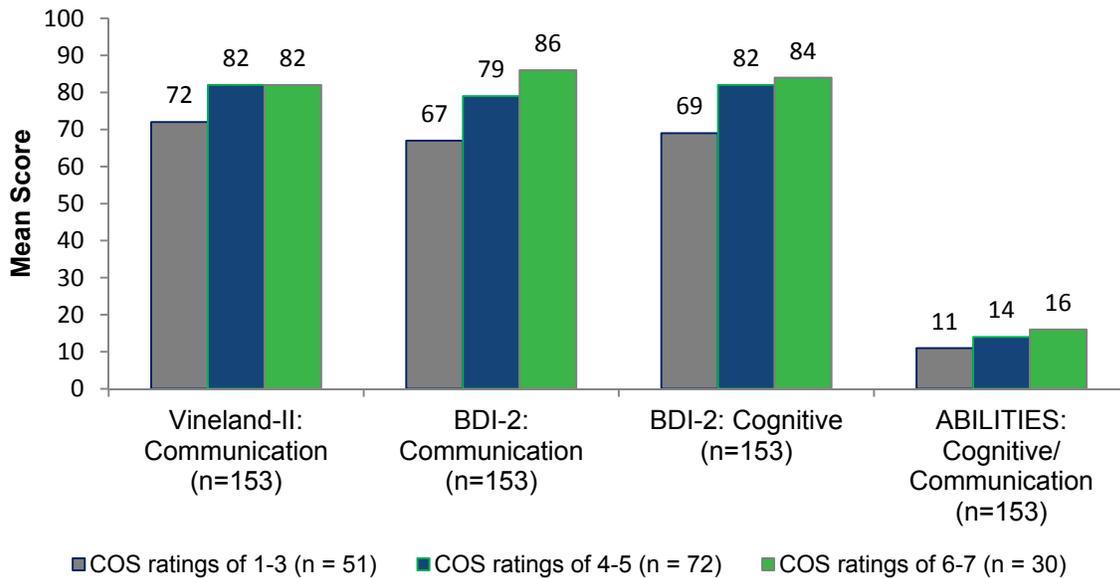
Action to Meet Needs

Across the assessment tools, the expected stair-step pattern generally was found for the three levels of COS ratings (1-3, 4-5, and 6-7) on the action to meet needs outcome. Higher mean scores were found for children with higher COS ratings (Exhibit 4.21).

- All of the differences between neighboring values reached a statistically significant threshold ($p < .05$), except for three of the comparisons involving the smallest group of children, the 35 children[15] who received ratings of 6 or 7 at entry.
  - Differences between children with COS ratings of 4-5 versus 6-7 showed the expected stair-step pattern, but did not reach statistical significance for the Vineland-II Motor subdomain or the BDI-2 Motor subdomain. Notably, the COS has been hypothesized to have variable relationships to assessment tools on the Motor subdomain, because children demonstrate various functional ways to get from place to place on the COS, without necessarily meeting the criteria to receive credit for motor skills on many assessment tools. Data for motor subdomains are presented

---

[15]   For action to meet needs, there were five comparisons involving this group of 35 children and three of the five did not reach statistical significance. Likewise, for knowledge and skills, there were four comparisons involving a group of 30 children and two of those did not reach statistical significance. The number of children in other subgroups being compared was larger (*n*'s ranged from 41-77) and 14 of the 15 comparisons reached statistically significant thresholds.

because states vary in how much and in what ways they rely on motor assessment information in outcomes measurement.

o Differences between children with COS ratings of 4-5 versus 6-7 showed the expected stair-step pattern, but were not strong enough to reach statistical significance for the Vineland-II Daily Living Skills subdomain; however the BDI-2 Adaptive subdomain did identify statistically significant differences between these two groups.

Exhibit 4.21   Mean Assessment Tool Scores by 3 Groups of COS Entry Ratings (1-3, 4-5, and 6-7) for Action to Meet Needs in Entry Sample (*n* = 153)
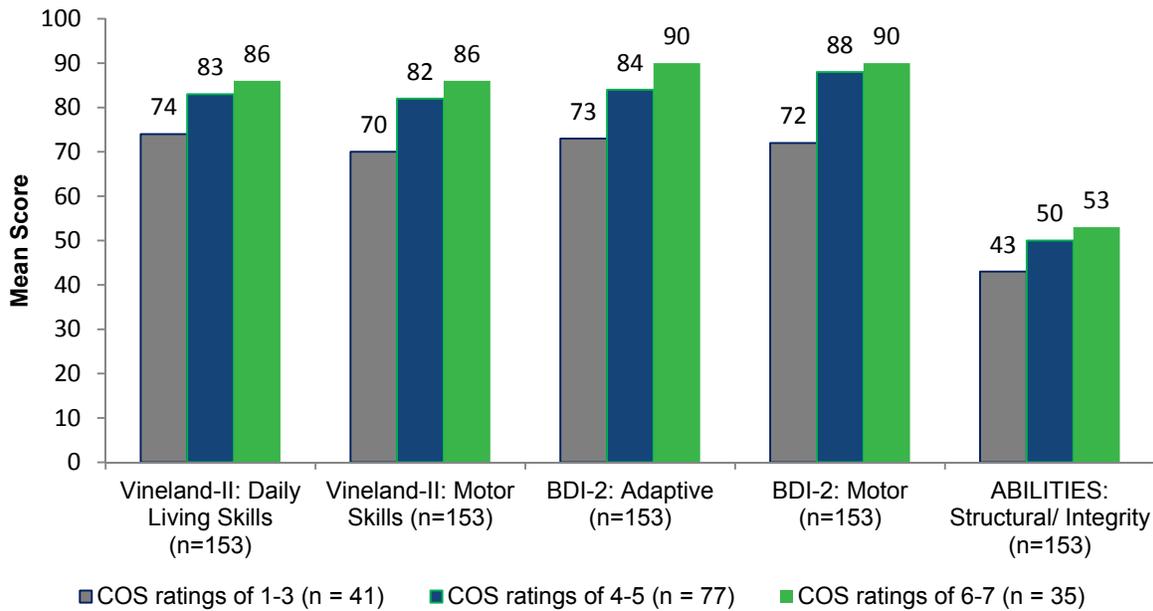
Despite the small sample size and the complexities of measuring domains versus outcomes, a consistent pattern of relationships was demonstrated that supports the validity of COS data. Children with lower COS ratings had lower average scores on the BDI-2, Vineland-II, and the reversed ABILITIES Index, and children with higher COS ratings had higher scores on these assessment tools.

Correlations between the COS and assessment tools at entry

Methodological Note about Correlations

A traditional way to investigate the relationships between approaches is to look at the correlations between COS ratings and scores on the various subdomains of standard assessment instruments, such as the BDI-2 and Vineland-II. Next, we present findings from data involving correlations, including using a multi-trait multi-method matrix (MTMM). As correlational findings of the COS with other assessment instruments are reviewed, the following are important caveats to consider:

1. **Truncated Variance at the Top End of the Distribution of Children.** The COS is designed to provide information about functioning for young children with disabilities participating in EI or ECSE services. As such, ratings are designed to make distinctions only at the lowest end of the scale. That is, a large portion of the distribution of ability in the population of young children birth through five years, where there is considerable variation in assessment tools scores, is only represented by ratings of 6 and 7 on the COS rating scale. When examining correlations between assessment tool scores and COS ratings, the restricted range at the top end of the COS can attenuate the observed correlation (Cohen, Cohen, West, & Aiken, 2003).

2. **Interval Scale Assumptions.** The BDI-2 and Vineland-II used standardized scores that meet interval scale requirements. However, the COS was designed as an ordinal scale. The 7 points on the COS represent distinct ordinal categories of functioning; there is no expectation or claim that there are equal distances between the rating points. Pearson correlations assume the presence of an interval scale. The current correlations included both interval-scaled variables (variables based on the BDI-2 and Vineland-II) and ordinal-scaled variables (COS ratings). Pearson correlations are presented here despite some violation of standard assumptions.

3. **Overlap Between Constructs Measured.** In interpreting correlations within a MTMM, the expectation is that correlations between measures of the same construct will be higher than correlations between measures of different constructs. This assumption is challenged by two features in the case of the COS:

   - "The" construct that is presumed to be the same is not a single construct. The BDI-2 and Vineland-II measure developmental domains, whereas the COS measures three functional child outcomes reflecting skills and behaviors from multiple domains. As a result, correlations between scores matched on the "same" construct (e.g., BDI-2 Cognitive; BDI-2 Communication; Vineland-II Communication; COS- Knowledge and Skills) are expected to be attenuated relative to correlations between scores from different tools that are measuring a single construct.

   - Development is highly integrated across domains in infants and young children, resulting in less differentiation or discrimination in functioning between constructs (whether domains or outcomes). Foundational skills (e.g., emerging language, motor control, or attention) in one domain are often necessary for children to build skills and/or to demonstrate their skills in other domains. As a result, scores for young children across domains tend to be more similar than often observed in youth or adult populations.

The combination of these features means that correlations across the "same" construct may be lower than expected and correlations across different constructs may be higher than expected in traditional MTMM analyses conducted on other populations and with different constructs. That is, the combination of these features means that the traditional pattern of high correlations between two measures of the same construct (e.g., two measures of Domain 1 or Trait 1) and low or no correlations of that domain with other, different constructs (e.g., any measure of Domain 1 with Domain 2 or any measure of Trait 1 with Trait 2), is less likely to be observed.

Overview of Correlational Results

Exhibit 4.22 provides a summary of correlations of each of the assessment tools and the COS at entry and between the BDI-2 and Vineland-II at entry. Correlations are presented with and without the motor subdomains; however, in this case, the use of the motor scores makes minimal difference in the correlations. From these data, we found that:

- The COS ratings are moderately correlated with each other across the outcomes (.67-.70).
- Correlations between the COS and the BDI-2 (.30-.46) and the COS and the Vineland-II (.28-.43) are only slightly weaker than those observed between the BDI-2 and the Vineland-II (.42-.66). This is true despite the following:
  - The BDI-2 and Vineland-II both measure domains, whereas the COS measures outcomes and,
  - The BDI-2 and Vineland-II assessments were usually conducted on the same day by the same assessor, whereas the COS was often completed 3-4 weeks prior to the assessments using a team process.
- Similar findings emerge when motor subdomains are included in the standardized assessments. Then, the correlations were more similar. The correlations between the COS and the BDI-2 (.30-.46) and the COS and the Vineland-II (.28-.50) are roughly comparable to those observed between the BDI-2 and the Vineland-II (.36-.74).

Exhibit 4.22   Overview of Range of Entry Correlations between Groups by Tool in Entry Sample (*n* = 153)

Relationships Examined	Range of Correlations Across the Three Outcome Areas at Entry	
	Not including Motor on BDI-2 and Vineland-II	Including Motor on BDI-2 and Vineland-II
COS ratings of 3 outcome areas with each other	.67-.70	
BDI-2 scores in subdomains with each other	.55-.78	.50-.78
Vineland-II scores in subdomains with each other	.43-.59	.43-.59
COS ratings with BDI-2 scores	.30-.46	.30-.46
COS ratings with Vineland-II scores	.28-.43	.28-.50
BDI-2 scores with Vineland-II Scores.	.42-.66	.36-.74

Multi-Trait, Multi-Method Matrix: Expected Relationships

A multi-trait, multi-method matrix (MTMM) is one way to examine relationships across assessment results. This approach has been used in studies of construct validity, especially when multiple methods are expected to be measuring the same construct(Campbell & Fiske, 1959). Researchers examine different sets of correlations, including correlations between measurements of the same construct using different data collection approaches or sources (e.g. interview, observation, direct assessment). We apply this approach to the study of COS validity by describing the expected patterns, how they relate to the existing data, findings observed, and interpretations for each. Data discussed are found in Exhibit 4.23.

MTMM Expected Pattern A. Looking at the same construct and same approach, there should be strong correlations.

Relationship to Study 2 Data. When there are multiple measures of the construct within an approach, there should be strong correlations.

Findings Observed

- Within the COS, this could not be examined because there is only one rating for each outcome using the COS process.
- Within the BDI-2:
  o There is only one score for positive social relationships, so it could not be tested.
  o For knowledge and skills, the correlation between cognitive and communication was .68.
  o For action to meet needs, the correlation between adaptive and motor subdomains was .60.
  o These correlations (.60 and .68) are stronger than the correlations of most of the other the BDI-2 subdomains with each other.
    ▪ BDI-2 correlations for knowledge and skills subdomains with action to meet needs subdomains were .50, .57, .58, and .55.
    ▪ BDI-2 correlations for action to meet needs subdomains with positive social relationships subdomains were .55 and .67.
    ▪ BDI-2 correlations for knowledge and skills subdomains with social relationships subdomain were .69 and .78.
    ▪ Overall 5 of the 8 correlations were consistent with the expected pattern (63%).
- Within the Vineland-II:
  o There is only one score for positive social relationships and knowledge and skills, so it could not be tested.
  o For action to meet needs, the correlation between the Daily Living Skills and Motor domain was .58.
  o This correlation ($r = .58$) is stronger than most of the other correlations of the Vineland-II subdomains with each other:

- Vineland-II correlations for action to meet needs with knowledge and skills were .43 and .45.
- Vineland-II correlations for action to meet needs with positive social relationships were .55 and .59.
- Vineland-II correlation for positive social relationships and knowledge and skills was .53.
- Overall, 4 out of the 5 correlations are consistent (80%).

Interpretation.

ENHANCE collected relatively few data that measured the same construct using the same approach. There were no instances for the COS, two for the BDI-2, and only one for the Vineland-II. Correlations within approach in general were stronger within the BDI-2 than within the Vineland-II, but for both tools, the correlations between two measures of the same construct measured using the same approach were among the strongest correlations observed for that approach. Overall, these findings supported the validity of the COS process.

MTMM Expected Pattern B. Looking at the same construct with different assessment tool approaches, the correlations should be stronger than looking at the different constructs with different assessment tool approaches.

Relationship to Study 2 Data. Across assessment tool approaches, certain subdomains were expected to map on to certain outcomes (or constructs). The correlations for the subdomains expected to map onto the outcomes should be higher than correlations between the subdomains not expected to map on to that outcome.

Findings Observed
- For the COS and BDI-2:
  - For positive social relationships, the correlation between the expected subdomain on the BDI-2 and the COS rating was .46. This was higher than correlations of all of the four BDI-2 subdomains not expected to map as strongly onto positive social relationships (r's = .34, .43, .35, .46).
  - For knowledge and skills, the correlations between the expected subdomains on the BDI-2 and the COS rating were .46 and .41. These were higher than correlations of all of the three BDI-2 subdomains not expected to map as strongly onto knowledge and skills (r's = .43, .37, .41).
  - For action to meet needs, the correlations between the expected subdomains on the BDI-2 and the COS rating were .46 and .46. These were higher than correlations for all of the three BDI-2 subdomains not expected to map as strongly onto action to meet needs (r's = .38, .30, .35).
  - Overall, all of the correlations (100%) across all of the constructs for the COS and BDI-2 produced the expected pattern. However, the size of differences in correlations between the subdomains mapped to each outcome and those that were not was not very large and, in most cases, did not reach statistical significance.

- For the COS and Vineland-II:
  - For positive social relationships, the correlation between the expected subdomain on the Vineland-II and the COS rating was .28. This was lower than correlations of all of the three Vineland-II subdomains not expected to map as strongly onto positive social relationships (r's = .32, .43, .33).
  - For knowledge and skills, the correlation between the expected subdomain on the Vineland-II and the COS rating was .34. This was higher than correlations of all of the three Vineland-II subdomains not expected to map as strongly onto knowledge and skills (r's = .29, .32, .30).
  - For action to meet needs, the correlations between the expected subdomains on the Vineland-II and the COS rating were .40 and .50. These were higher than correlations of both of the Vineland-II subdomains not expected to map as strongly onto action to meet needs (r's = .31 and .32).
  - For two of the three constructs, (knowledge and skills and action to meet needs) the Vineland-II and COS correlations produced the expected pattern. As with the BDI-2, the size of differences in correlations between the subdomains that were mapped to each outcome and those that were not was not very large and, in most cases, did not reach a statistical significance.
  - For positive social relationships, none of the correlations was consistent with the expected pattern. It is possible that the child's scores on functioning in positive social relationships were influenced by the parent caregiver responses across all areas in a stronger way. That is, for the Vineland-II, information about each of the subdomains is taken from caregivers' reports of the child's functioning in daily routines, whereas in the BDI-2, the child's scores are influenced more by direct assessment.
- For the BDI-2 and Vineland:
  - For positive social relationships, the correlation between the expected subdomain on the BDI-2 and the Vineland-II was .61. This was higher than correlations of all of the three subdomains not expected to map as strongly onto positive social relationships (r's = .55, .55, .48).
  - For knowledge and skills, the correlations between the expected subdomains on the BDI-2 and the Vineland-II were .52 and .51. This was the same as or higher than correlations of all of the other subdomains not expected to map as strongly onto knowledge and skills (r's = .47, .52, .46, .42, .36, and .44).
  - For action to meet needs, the correlation between the expected subdomains on the BDI-2 and the Vineland-II were .55, .66, .74, and .52. These were higher than nearly all of the correlations of subdomains not expected to map as strongly onto action to meet needs (r's = .48, .51, .54, and .50).
  - Overall, the pattern of correlations across all of the constructs for the Vineland-II and BDI-2 produced the expected pattern. Correlations were slightly higher than observed for the COS with the BDI-2 and the COS with the Vineland-II. The size of the distinctions in correlations between the subdomains that were mapped to each outcome and those that were not, was not very large, and in most cases, did not reach a statistically significant threshold.

Interpretation.

Across all the different assessment tool approaches, correlations between subdomains expected to map to the same construct generally were higher across assessment tools than correlations that were not expected to map to the same construct. The only major exception to this was found with the COS and Vineland-II correlations for positive social relationships. Correlations were slightly higher between Vineland-II and BDI-2 than between COS and either of the other tools. This pattern was expected because the BDI-2 and Vineland-II both measure domains and were conducted at the same point in time.

As found with the COS correlations with the BDI-2 and Vineland-II, the size of differences in correlations between the subdomains that were mapped to each outcome and those that were not, was not very large and, in most cases, did not reach a statistical significance. This finding is consistent with the expectations outlined earlier about patterns in the correlations.

MTMM Expected Pattern C. Looking at different constructs with same assessment tool approaches, the correlations should be weaker than looking at the same constructs with different assessment tool approaches.

- **Relationship to Study 2 Data**. Correlations across the outcomes within an assessment tool approach (within COS, within BDI-2, within Vineland-II) should be weaker than correlations that map to the same outcome but are measured with different assessment tools.[16]

Findings Observed

- For the COS:
  - For positive social relationships, correlations across outcomes within the same assessment approach were .70 and .67. Correlations for the subdomains mapped to the same outcome with different assessment approaches were .46 and .28.
  - For knowledge and skills, correlations across outcomes within the same assessment approach were .70 and .67. Correlations for the subdomains mapped to the same outcome with different assessment approaches were .46, .41, and .34.
  - For action to meet needs, correlations across outcomes within the same assessment approach were .67 and .67. Correlations for the subdomains mapped to the same outcome with different assessment approaches were .46, .46, .40, and .50.
- For the BDI-2:
  - For positive social relationships, correlations across outcomes within the same assessment approach were .69, .78, .55, and .67. Correlations for the subdomains mapped to the same outcome with different assessment approaches were .46 and .28.

---

[16] Note that while this is the expected pattern in an MTMM, under methodological consideration, we outlined several reasons why the assessment tool and/or approach in this study is likely to have a more pronounced influence on the strength of the relationship than in most MTMMs and why the correlations across outcomes are likely to be stronger in this early childhood population.

- o For knowledge and skills, correlations across outcomes within the same assessment approach were .69, .78, .50, .58, .57, and .55. Correlations for the subdomains mapped to the same outcome with different assessment approaches were .46, .41, and .34.
  - o For action to meet needs, correlations across outcomes within the same assessment approach were .55, .67, .50, .57, .58, and .55. Correlations for the subdomains mapped to the same outcome with different assessment approaches were .46, .46, .40, and .50.
- For the Vineland-II:
  - o For positive social relationships, correlations across outcomes within the same assessment approach were .53, .59, and 55. Correlations for the subdomains mapped to the same outcome with different assessment approaches were .46 and .28.
  - o For knowledge and skills, correlations across outcomes within the same assessment approach were .43, .45, and .53. Correlations for the subdomains mapped to the same outcome with different assessment approaches were .46, .41, and .34.
  - o For action to meet needs, correlations across outcomes within the same assessment approach were .59, .55, .43, and .45. Correlations for the subdomains mapped to the same outcome with different assessment approaches were .46, .46, .40, and .50.

Interpretation.

Overall, the original MTMM expected pattern was not at all supported for any of the outcomes on the COS or BDI-2 and showed only minimal and mixed support on the Vineland. We suggest that these findings are due to the fact that the BDI-2 and Vineland-II measure domains, whereas the COS measures outcomes, and from the stronger correlations across outcomes common for developmental domains for young children.

MTMM Expected Pattern D. Correlations between subdomains not expected to load on the same construct across different assessment tool approaches will be the lowest observed correlations. They will be lower than correlations between subdomains that are expected to load on the same construct even though they are from different assessment tool approaches and lower than subdomains not expected to load on the same construct but using the same assessment tool approach.

- **Relationship to Study 2 Data**. Correlations for subdomains not expected to map onto a given outcome/construct across assessment tool approaches will be lower than both the correlations across assessment tool approaches that are expected to load on the outcome and the correlations across outcomes that share the same assessment tool approach.

Findings Observed

- The off-diagonal cells in the portion of the MTMM matrix providing correlations across assessment tools are expected to be the lowest correlations observed. For positive social relationships between the COS and the BDI-2 and Vineland-II, these were .34, .43, .35, .46, .32, .43, and .33.

o In contrast, the diagonals (where the subdomains are mapped to the specific outcome/construct) were .46 and .28 (showing some support for the BDI-2 but not for the Vineland-II).

o In contrast, the correlations across outcomes from a single assessment tool method range from .67-.70 for COS, .50-.78 for BDI-2, and .43-.59 for Vineland-II (showing support for the COS and BDI-2 and some support for the Vineland-II).

- For knowledge and skills, between the COS and the BDI-2 and Vineland-II these correlations were .43, .37, .41, .29, .32, and .30.

  o In contrast, the diagonals (where the subdomains are mapped to the specific outcome/construct) were .46, .41, and .34.

  o In contrast, the correlations across outcomes from a single assessment tool method range from .67-.70 for COS, .50-.78 for BDI-2, and .43-.59 for Vineland-II.

- For action to meet needs, between the COS and the BDI-2 and Vineland-II these were .38, .30, .35, .31, and .32.

  o In contrast, the diagonals (where the subdomains are mapped to the specific outcome/construct) were .46, .46, .40, and .50 (Supported 100%).

  o In contrast, the correlations across outcomes from a single assessment tool method range from .67-.70 for COS, .50-.78 for BDI-2, and .43-.59 for Vineland-II.

Interpretation.

Consistent with MTMM expectations, the correlations for cells across different constructs and across different methods were generally the smallest. However, the discrepancy in the size of correlations was smaller than what is expected in other samples or MTMMs focusing on more unified constructs.

Testable Expected Pattern E[17]. Based on the definition of the three child outcomes, the communication subdomain should show moderate relationships across all three outcome areas, but should show the strongest relationship with the knowledge and skills outcome area.

- **Relationship to Study 2 Data**. Correlations for the BDI-2 Communication subdomain and the Vineland-II Communication subdomain should be moderate across all three outcomes. The correlations should be the strongest with the knowledge and skills outcome when compared to positive social relationships and action to meet needs.

---

[17] This expected pattern is not a routine part of relationships examined in most MTMM work, but is an outgrowth of the project's understanding of the three child outcomes.

Findings Observed

- BDI-2 Communication
  - The Communication subdomain was correlated .46 with knowledge and skills, .34 with positive social relationships, and .30 with action to meet needs.
- Vineland-II Communication
  - The Communication subdomain was correlated .34 with knowledge and skills, .32 with positive social relationships, and .32 with action to meet needs.

Interpretation.

For both the BDI-2 and the Vineland-II, communication was related to all three outcomes. Correlations for communication were within +/ - .10 of the correlation found for the subdomain most expected to map onto each outcome. For both the BDI-2 and the Vineland-II, the correlation of communication with knowledge and skills was higher than that found with positive social relationships or action to meet needs. The evidence supports the expected patterns.

Summary

The pattern of relationships found in the MTMM matrix generally follow the expected relationships after taking into account unique considerations based on the study sample and constructs being examined. The ratings that teams assigned to children showed consistent, but not strong, relationships to the expected subdomains on assessment tools that matched the same constructs.

Exhibit 4.23   MTMM Results at Program Entry in Entry Sample (*n* = 153)

Entry Sample (*n* = 153)	COS			BDI-2					Vineland-II			
	A1 Positive Social Relationships	A2 Knowledge and skills	A3 Action to Meet Needs	B1 Personal Social	B2.a Communication	B2.b Cognitive	B3.a Motor	B3.b Adaptive	C1 Socialization	C2 Communication	C3.a Daily Living Skills	C3.b Motor Skills
COS												
A1 Positive Social Relationships												
A2 Knowledge and skills	0.70											
A3 Action to Meet Needs	0.67	0.67										
BDI-2												
B1 Personal Social	0.46	0.43	0.38									
B2.a Communication	0.34	0.46	0.30	0.69								
B2.b Cognitive	0.43	0.41	0.35	0.78	0.68							
B3.a Motor	0.35	0.37	0.46	0.55	0.50	0.58						
B3.b Adaptive	0.46	0.41	0.46	0.67	0.57	0.55	0.60					
Vineland-II												
C1 Socialization	0.28	0.29	0.31	0.61	0.47	0.52	0.48	0.54				
C2 Communication	0.32	0.34	0.32	0.55	0.52	0.51	0.51	0.50	0.53			
C3.a Daily Living Skills	0.43	0.32	0.40	0.55	0.46	0.42	0.55	0.66	0.59	0.43		
C3.b Motor Skills	0.33	0.30	0.50	0.48	0.36	0.44	0.74	0.52	0.55	0.45	0.58	

Note: Boxes around correlations indicate the subdomains where content is expected to be most closely related to that in the specific outcome area (i.e., positive social relationships, knowledge and skills, or action to meet needs).

Correlations between COS ratings and assessment tool scores at exit

Exhibits 4.24 and 4.25 provide the same information as shown above in the entry sample, but these show the correlations and MTMM for the exit correlations in the longitudinal sample. For these analyses, the sample drops from *n* = 153 to *n* =70. These data provide information about the relationships between the assessment tools and COS ratings at exit. However, based on the sample size, they are not likely to be reliable population estimates.

- Correlations at exit show many of the same relationships as described for the entry sample.
- Exit correlations tended to be higher[18] than entry correlations for the COS, the BDI-2, and for the Vineland-II.
    - For the COS and BDI-2, 11 out of 15 correlations were higher at exit than entry. Three of the four anomalies where correlations involving the BDI-2 motor subdomain.
    - For the COS and Vineland-II, 11 out of 12 correlations were higher at exit than entry.
    - For the BDI-2 and Vineland-II, 15 out of 20 correlations were higher at exit than entry. Four motor correlations decreased at exit and one additional correlation remained the same at both time points.
- At exit, correlations across the three outcome areas are of similar strength (.82, .83, .80).
- Motor subdomain correlations were somewhat weaker at exit than those observed at entry. This is consistent with understanding that motor plays less of a role in allowing children to demonstrate understanding and skills by age 3 than it does among infants/toddlers and younger children.
- All three assessment tools have stronger correlations across outcomes within their own approach[19] at exit than they did at entry. The increase in the strength of correlations from entry to exit was fairly consistent across the three assessment tools.

---

[18] Note that exit correlations were based on the subset of children in the longitudinal sample (*n* = 70), whereas entry correlations were based on children in the entry sample *(n* = 153); the longitudinal sample is a subset of the entry sample.

[19] Note that entry and exit assessments were conducted too far apart to interpret these correlations within approach as short-term test-retest reliability.

Exhibit 4.24   Overview of the Range of Exit Correlations between Groups by Assessment Tool in the Longitudinal Sample (*n* = 70)

Relationships Examined	Range of Correlations Across the Three Outcome Areas at Exit	
	Does not include Motor on BDI- 2 and Vineland-II	Includes Motor on BDI-2 and Vineland-II
COS ratings in 3 outcomes areas with each other	.80-.83	
BDI scores in subdomains with each other	.47-.72	.26-.72
Vineland-II scores in subdomains with each other	.69-.74	.55-.74
COS ratings with BDI scores	.41-.62	.28-.62
COS ratings with Vineland-II scores	.40-.62	.40-.62
BDI scores with Vineland Scores	.46-.77	.39-.77

Exhibit 4.25   MTMM Results at Program Exit in Longitudinal Sample (*n* = 70)

Longitudinal Sample (*n* = 70)		COS			BDI-2					Vineland-II			
		A1 Positive Social Relation-ships	A2 Knowledge and skills	A3 Action to Meet Needs	B1 Personal Social	B2.a Communi-cation	B2.b Cognitive	B3.a Motor	B3.b Adaptive	C1 Sociali-zation	C2 Communi-cation	C3.a Daily Living Skills	C3.b Motor Skills
COS	A1 Positive Social Relationships												
	A2 Knowledge and skills	0.82											
	A3 Action to Meet Needs	0.83	0.80										
BDI-2	B1 Personal Social	0.62	0.49	0.58									
	B2.a Communication	0.55	0.58	0.62	0.65								
	B2.b Cognitive	0.52	0.45	0.53	0.60	0.72							
	B3.a Motor	0.29	0.28	0.36	0.26	0.24	0.36						
	B3.b Adaptive	0.41	0.44	0.47	0.62	0.56	0.47	0.38					
Vine-land-II	C1 Socialization	0.54	0.51	0.56	0.72	0.59	0.46	0.41	0.64				
	C2 Communication	0.57	0.62	0.58	0.60	0.65	0.63	0.39	0.60	0.69			
	C3.a Daily Living Skills	0.40	0.44	0.41	0.62	0.55	0.48	0.48	0.77	0.73	0.74		
	C3.b Motor Skills	0.44	0.43	0.53	0.53	0.49	0.46	0.45	0.52	0.51	0.55	0.60	

Note: Boxes around correlations indicate the subdomains where content is expected to be most closely related to that in the specific outcome area (i.e., positive social relationships, knowledge and skills, or action to meet needs).

Relationships of COS ratings and assessment scores across outcomes over time

Exhibit 4.26 shows the relationships between COS ratings on the three outcomes areas at entry (upper left portion of exhibit), at exit (lower right portion of exhibit), and between entry and exit (lower left portion of exhibit). These data are based on the longitudinal sample (*n* = 70).

- Correlations between the outcomes based on COS ratings were higher at exit than at entry.
- Correlations between entry to exit were of similar strength across all three outcome areas.

Exhibit 4.26 Correlations between Entry and Exit COS Ratings in Longitudinal Sample (*n* = 70)

	Entry-Positive Social Relation-ships	Entry-Know-ledge and Skills	Entry-Action to Meet Needs		Exit-Positive Social Relation-ships	Exit-Know-ledge and Skills	Exit-Action to Meet Needs
Entry-Positive Social Relationships							
Entry-Knowledge and Skills	0.70						
Entry-Action to Meet Needs	0.77	0.68					
Exit-Positive Social Relationships	0.46	0.41	0.41				
Exit-Knowledge and Skills	0.42	0.47	0.39		0.82		
Exit-Action to Meet Needs	0.42	0.48	0.47		0.83	0.80	

Relationships in Determinations of Progress between COS and Assessment Tools

Earlier in this section, we described the distributions of change observed in the COS ratings data. However, states do not submit specific ratings information for accountability reporting. Rather, this information is aggregated into the percentages of children in each of five progress categories that are then numerically converted into two summary statement percentages. States set targets around summary statement percentages and are monitored for changes in those across years (see additional information in Section 1 and Appendix C). Below we examine how data from the COS corresponds to what could be derived using other approaches to derive the summary statement results.

Correspondence between progress categories on COS, BDI-2, and Vineland-II

When the COS was established, guidance was provided about how to use information from the ratings to determine progress categories for federal reporting. However, neither the BDI-2, nor the Vineland-II was designed to produce scores that translate to progress categories and no specific guidance is provided by the publishers about one preferred, validated approach for conversion.[20] Earlier in this section (Data Analysis), we described the methods we used to map the BDI-2 and Vineland-II scores to progress categories. While we describe correspondence information in this section, there is no "gold standard" mapping assessment tool scores on to one progress category classifications.

Exhibit 4.27 describes the correspondence between using COS ratings, BDI-2 scores, and Vineland-II scores to determine progress categories for the 70 children in the longitudinal sample. Exhibit 4.28 shows definitions for each progress category. For each combination of approaches, information is provided about the percent of exact progress category agreement overall and a Kappa for agreement across all five categories. Kappa values also are provided for agreement on each specific progress category (i.e., examining progress category e versus not e for instance rather than all five categories together). Kappa values range from -1.0 to 1.0, with 1.0 indicating perfect agreement. Landis and Koch (1977) provide guidelines for interpreting kappa values, with values from 0.0 to 0.2 indicating slight agreement, 0.21 to 0.40 indicating fair agreement, 0.41 to 0.60 indicating moderate agreement, 0.61 to 0.80 indicating substantial agreement, and 0.81 to 1.0 indicating almost perfect or perfect agreement.

The findings show:

- Substantial agreement was not observed in correspondence across any of the approaches, indicating that various assessment tools either reflect somewhat different perspectives on the skills of different children and/or that algorithms to convert scores or ratings into progress categories have important differences in how they classify children.

- Most overall Kappas indicate slight agreement (75% of Kappas shown), with fair agreement found between the COS and both BDI-2 and Vineland-II for positive social relationships.

- Overall Kappa values are slightly higher between the BDI-2 and Vineland-II, but still are predominantly in the fair range (80% of Kappas shown). Correspondence may be influenced by conversions from standard scales involving similar decision rules, from both assessment tools measuring children on domains, and/or from assessing the child on the same day rather than after some time delay.

- Agreement tends to be the highest for category e when looking at the COS with either of the assessment tools. The pattern shifts a bit, with somewhat stronger Kappa agreement for progress categories b and e.

---

[20] In fact, a subset of states who use assessments cores from the BDI-2 for federal reporting use a number of different decision rules for converting BDI-2 data into progress categories. There is not one consistent approach even among this group of states.

Exhibit 4.27 Percent Agreement and Kappas between COS and BDI-2 on Progress Categories Used for Federal Reporting

Comparison Longitudinal Sample (*n* = 70)	Exact Progress Category Agreement *n* (%)	Kappa for agreement across all categories	Kappa for specific progress category only				
			a**	b	c	d	e
COS and BDI-2							
Positive Social Relationships	33 (47%)	.27	-.01	.28	.03	.34	.33
Knowledge and Skills	20 (29%)	.08	-.01	.15	.02	.08	.06
Action to Meet Needs without Motor	21 (30%)	.06	-.02	-.07	-.09	.12	.22
Action to Meet Needs with Motor	22 (31%)	.11	.00	.08	.00	.13	.26
COS and Vineland-II							
Positive Social Relationships	29 (41%)	.21	-.02	.13	.13	.25	.33
Knowledge and Skills	24 (34%)	.14	-.02	-.04	.15	.13	.30
Action to Meet Needs	25 (36%)	.17	.00	.21	.07	.13	.29
Action to Meet Needs with Motor	24 (34%)	.12	-.01	-.02	-.00	.12	.33
BDI-2 and Vineland-II							
Positive Social Relationships	17 (63%)	.39	.66	.37	.52	.29	.41
Knowledge and Skills	11 (52%)	.19	.66	.25	-.04	.10	.35
Action to Meet Needs	39 (56%)	.38	No value*	.50	.33	.11	.46
Action to Meet Needs with Motor	22 (61%)	.44	.49	.53	.10	.40	.48

* No value was produced for kappa because neither the BDI nor the Vineland classified any children into progress category "a."

** Small numbers of children are in progress category a influencing Kappa rates for this category.

Exhibit 4.28   Letter Correspondence for Progress Categories Used in Federal Reporting

> (a) Did not make any progress
> (b) Made progress but not sufficient to move closer to same-age peers
> (c) Made progress and moved closer to same-age peers
> (d) Achieved functioning comparable to same-age peers
> (e) Maintained functioning comparable to same-age peers

Distributions of Progress Categories

Exhibits 4.29–4.32 show the number of children assigned to each progress category for the COS, BDI-2, and Vineland-II for the 70 children in the longitudinal sample, with data shown for each of the three child outcomes. The five progress categories add up to 100%, so higher percentages in one category will mean lower percentages in another.

- Different approaches had the highest percentage of children in progress category e across the three outcomes. The BDI-2 was highest for positive social relationships, the Vineland-II was highest for knowledge and skills, and the COS was highest for action to meet needs.

Exhibit 4.29   Distribution of Progress Categories for Positive Social Relationships as Identified by Different Approaches in the Longitudinal Sample (*n* = 70)
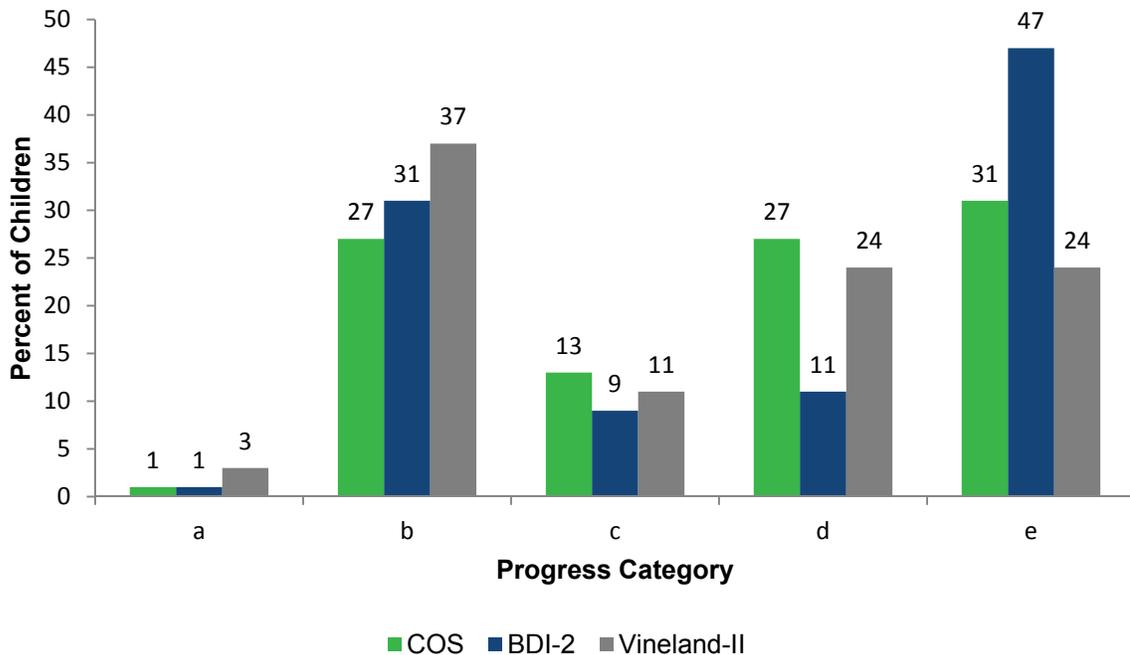
Exhibit 4.30   Distribution of Progress Categories for Knowledge and Skills as Identified by Different Approaches in the Longitudinal Sample (*n* = 70)
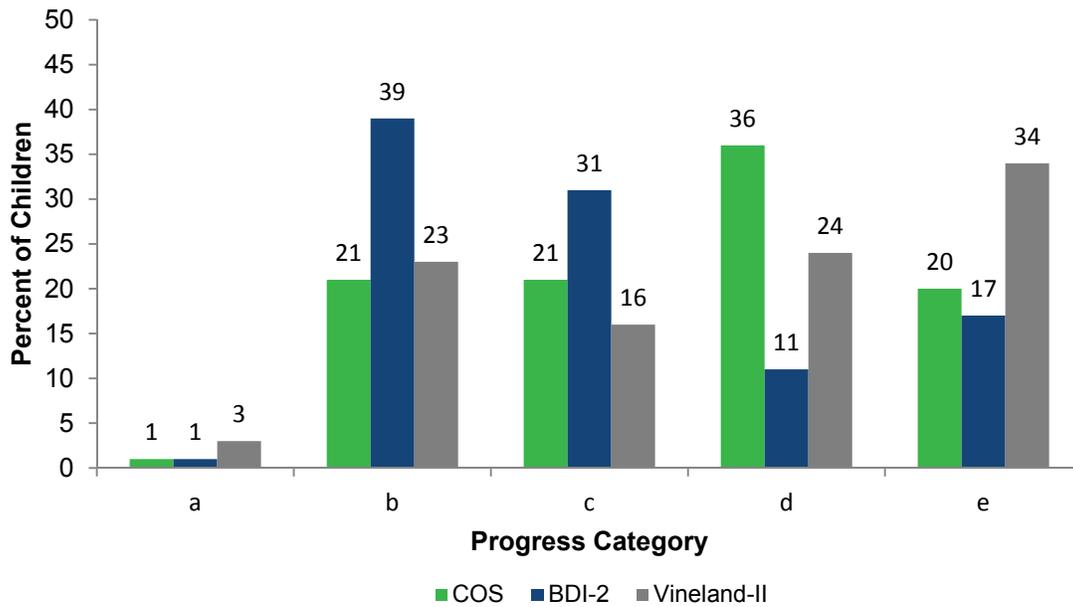
Exhibits 4.31 and 4.32 both show the distribution for progress categories by assessment tools for action to meet needs. Exhibit 4.31 shows the distribution when adaptive, daily living skills, and motor scores are considered; Exhibit 4.32 bases progress categories only on information from adaptive and daily living skills subdomains. These data indicate that a higher percentage of children are in category e when motor scores are not included for both assessment tools (Exhibit 4.32)

Exhibit 4.31   Distribution of Progress Categories for Action to Meet Needs as Identified by Different Approaches in the Longitudinal Sample (*n* = 70)
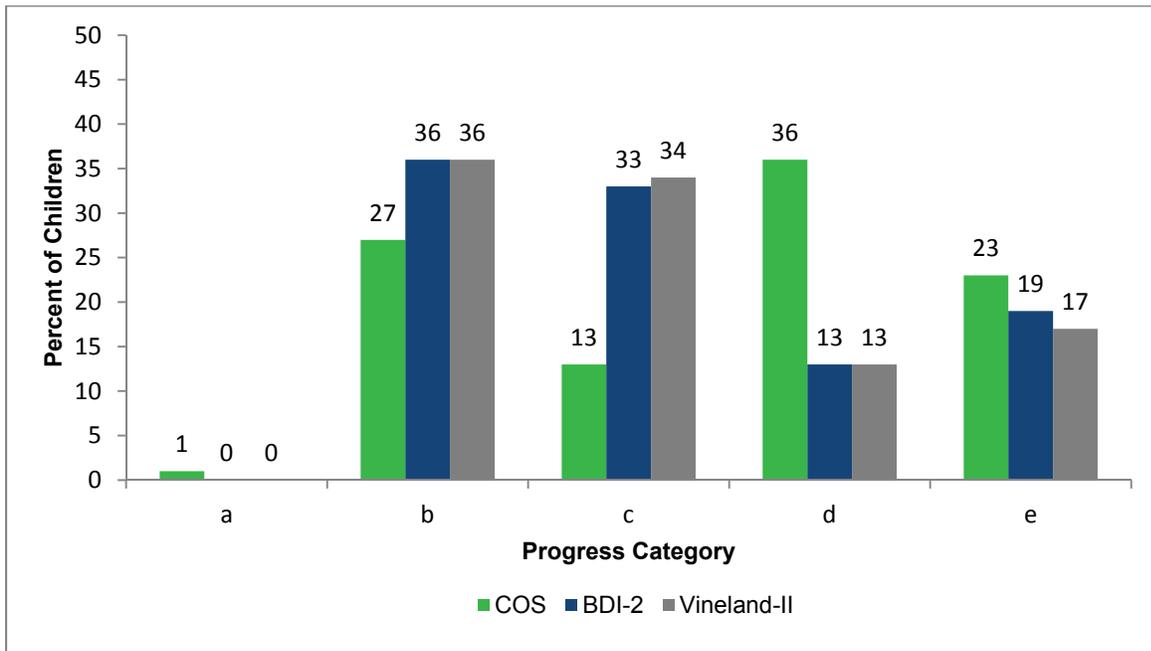
Exhibit 4.32   Distribution of Progress Categories for Action to Meet Needs Without Using Motor Subdomains as Identified by Different Approaches in the Longitudinal Sample (*n* = 70)
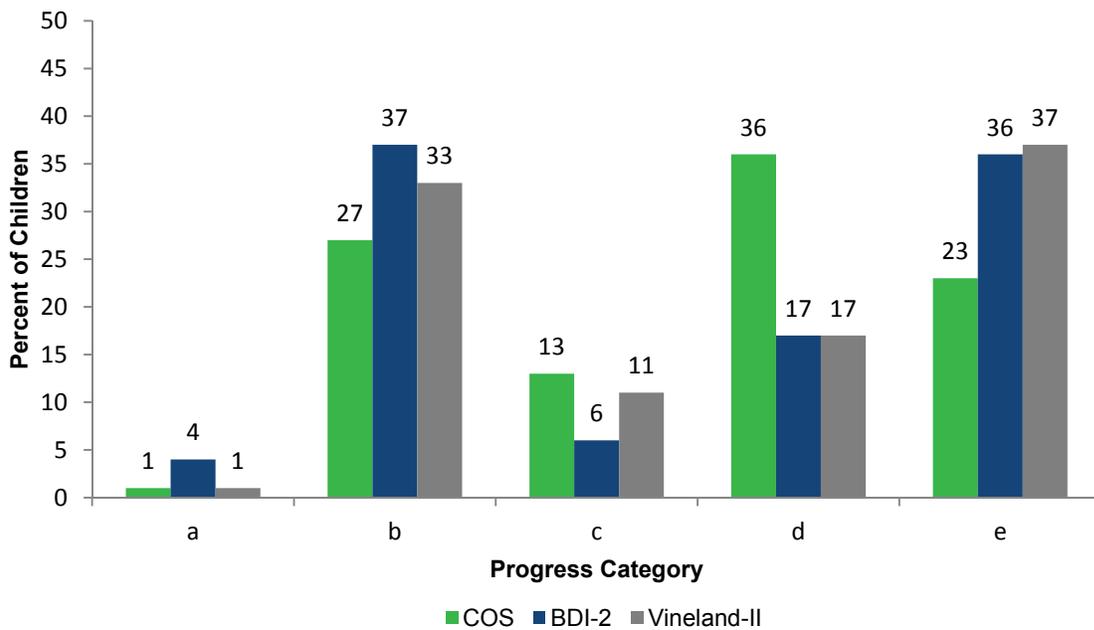
Characteristics of cases where classification disagreements occur

We analyzed data to determine if there were systematic relationships between key variables and disagreements in progress category classifications. No strong relationships were identified. For instance, when comparing those where there was classification agreement versus disagreement on knowledge and skills between the COS and the BDI-2, there were not significant differences in child age at entry, ABILITIES Index scores, number on team at entry or exit, parent on team at entry or exit, or service coordinator present on team. Also, classification disagreements did not predominate just in just one single program or state, but rather was observed across the sample.

Considering threshold for age expectations as a factor in progress category agreement

The main distinction between progress category b and e for many children is whether or not the child enters at or above age expectations.[21] Based on the variation in percentages of progress categories, it appears that different approaches tend to employ different thresholds to categorize children as at age-expected or not. While the cutoffs described in the data analysis section are one logical approach to determining where the line for age-expected functioning lies, other approaches could also be used. We undertook one such analysis to try to determine alternate cutoffs based on empirical data rather than using a conceptual basis.

To identify the optimal cutoff for age expectations for each of the standardized assessments, we identified the standard score that best differentiated or predicted children in the entry sample of the child assessments study who were rated as above and below age expectations on the COS. The first step in this analysis was to compute the area under the curve (AUC) for the Receiver Operating Curves (ROC) analysis for each assessment domain that was associated with each outcome area. We dichotomized the COS into 0 (below age expectations) or 1 (at or above age expectations) and then the standard score on the assessment tool was used to predict if a child scored above or below age expectations on the COS. Results showed poor prediction of age expectations on the COS using the standard scores. These findings were similar for both assessment tools and across all three outcome areas. Results showed that the area under the curve (AUC) was .7 or below for all comparisons using both the BDI-2 and the Vineland-II. This AUC value is considered a poor association for ROC criteria(Tape, n.d.). An AUC of .5 represents that the classification decision based on the prediction model is as likely to be wrong as it is to be right; it is a chance prediction equivalent to that of flipping a coin randomly. An AUC of 1.0 represents perfect prediction with no false positives. Therefore, an AUC value of .6-.7 is better than flipping a coin, but not too far from that age-old family favorite. Given these findings, we

---

[21] The developmental trajectories for children in these two progress categories are often very similar, the key distinction is in the child's level of functioning at entry, that is, whether or not the child is at age-expected functioning at the outset.

did not move forward with the next stage of analysis to identify an optimal cut point for age-expected functioning.

These findings are consistent with the contention that the BDI and Vineland are measuring a construct that is distinct from the outcomes measured by the COS and/or that teams are using multiple sources of information that go beyond assessment tool scores to make their appraisals of age-expected functioning.

Correspondence between summary statements based on COS, BDI-2, and Vineland-II

As outlined in the background section (Section 1) and shown in Appendix C, summary statement percentages are derived directly from the percentages of children in each of the five progress categories using a straightforward mathematical computation. Although progress category information is used for federal reporting and provides a picture of the kinds of progress children in EI and ECSE make, it is the summary statements for which states set target percentages and on which the U.S. Department of Education monitors state results mostly closely.

Exhibits 4.33 and 4.34 show the distribution of percentages on the two summary statements for each outcome by approach. Data are based on the longitudinal sample in the child assessments study *(n = 70)*.

Although Summary Statement 1 data are drawn from the longitudinal sample of the child assessments study, the *n*'s for Summary Statement 1 differ by approach and outcome (Exhibit 3.33). Differences occur because standard practice for computing Summary Statement 1 requires leaving children identified in progress category e out of the equation. The number of children in progress category e varies by approach and by outcome resulting in different values for each bar.

For each approach, n's are the following shown in order of outcomes: COS: *n* = 48, 56, 54; BDI-2: *n* = 37, 58, 57; Vineland-II: *n* = 53, 46, 58 for positive social relationships, knowledge and skills, and action to meet needs respectively.

- Percentages on Summary Statement 1 are higher when based on the COS than for the BDI-2 and Vineland-II. Differences were greater between the COS and BDI2 than between the COS and Vineland-II.
- Distributions on summary statements across approaches are more similar than distributions of progress categories across approaches.

Exhibit 4.33   Summary Statement 1: Greater Than Expected Progress Percentages as Determined by COS, BDI-2, and Vineland-II in Longitudinal Sample
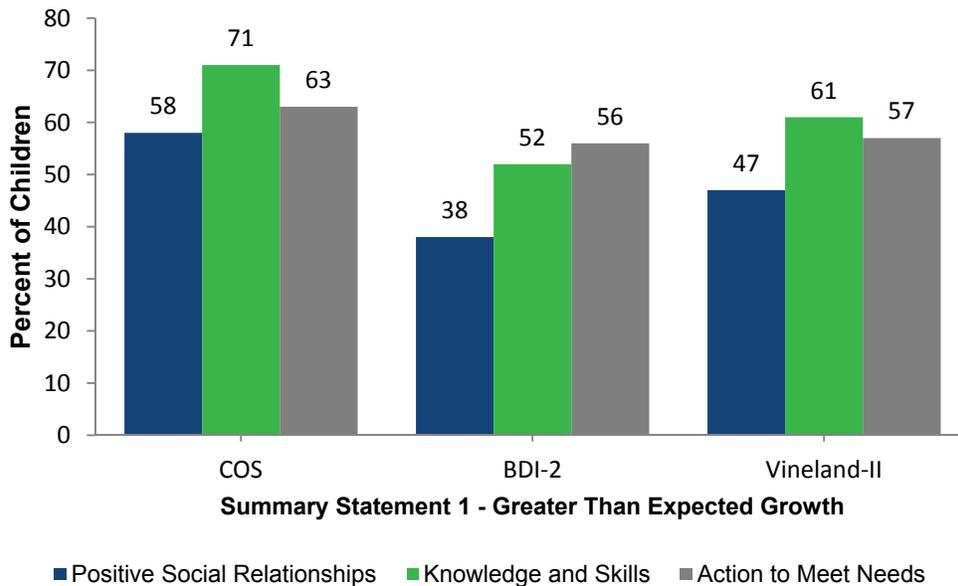
Summary Statement 2 percentages are based on the full longitudinal sample of 70 children for each tool (Exhibit 3.34).

- Percentages found for Summary Statement 2 are fairly consistent across the three outcomes on the COS.
- Percentages show much more variability for the BDI-2 and Vineland-II across outcomes.

Assessment tools that categorize higher numbers of children at age-expected levels will have higher percentages on Summary Statement 2. Children with similar developmental trajectories, but who start below the age-expected line often will end up in progress category b rather than e, and will not be reflected in Summary Statement 2 percentages.

Exhibit 4.34   Summary Statement 2: Exit at Age Expectations Percentages as Determined by COS, BDI-2, and Vineland-II on Longitudinal Sample (*n* = 70
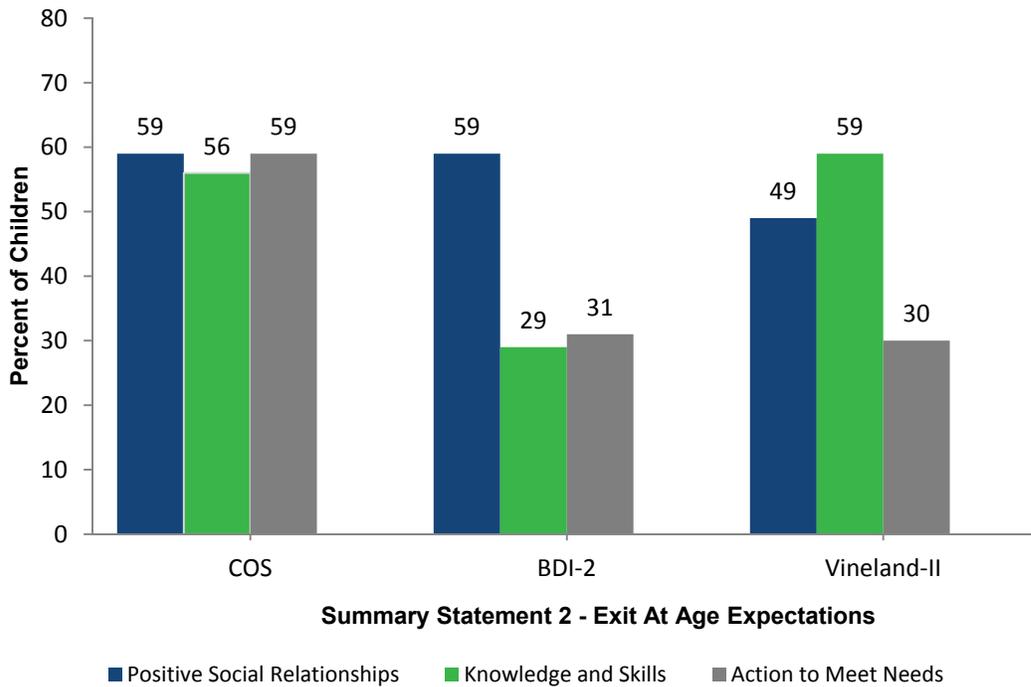
Examining children who change trajectories across COS, BDI-2, and Vineland-II

A follow-up question of interest was to understand what kind of change occurred on the assessment tools among the children that the COS identified as changing developmental trajectories in Summary Statement 1 (i.e., children in progress categories c and d who enter below age expectations and either catch up to age-expected levels or make greater than expected growth relative to peers). Exhibit 4.35 shows data for this subset of children and examines the change in COS ratings observed as well as changes in entry to exit assessment tool subdomain scores most relevant to the outcome area.

On average, those who the COS identified as changing trajectories:

- Showed statistically significant increases on COS ratings relative to those who were not trajectory changers on each of the three outcome areas.
- On average, showed more positive changes on both BDI-2 and Vineland-II subdomain scores relative to those who were not identified as trajectory changers. Despite the small sample size and wide standard deviation, some of these differences in the amount of positive change among trajectory changers or negative change in non-changers, reached a statistical significance.

Exhibit 4.35 Comparison of Changes in Mean COS Ratings, Mean BDI-2 Scores, and Mean Vineland-II Scores among Children who Change Trajectories and Do Not on the COS in Longitudinal Sample (possible *n* = 70)

	Positive Social Relationships		Knowledge and Skills		Action to Meet Needs	
	Not Trajectory Changers M (SD)	Trajectory Changers M (SD)	Not Trajectory Changers M (SD)	Trajectory Changers M (SD)	Not Trajectory Changers M (SD)	Trajectory Changers M (SD)
n (based on COS trajectory changers)	20	28	16	40	20	34
COS						
Entry Ratings	4.6 (1.8)	3.7 (1.3)	4.4 (1.6)*	3.4 (1.3)*	4.2 (1.5)	3.7 (1.3)
Exit Ratings	3.8 (1.4)	6 (1)***	3.9 (1.4)***	5.7 (1.1)***	3.7 (1.3)	6 (1)***
Change in COS	-.8 (.9)	2.3 (1.4)***	-.5 (.9)***	2.2 (1.2)***	-.5 (.9)	2.3 (1.3)***
BDI-2						
Subdomain(s)	Personal-Social		Cognitive / Communication		Adaptive / Motor	
Entry Scores	83.8 (14)	82.7 (11.2)	79.2 (18.5) / 77.5 (15.1)	81.4 (16.2) / 77.3 (17)	84.2 (14.5) / 81.8 (12.6)	81 (14.9) / 83.6 (15.7)
Exit Scores	78.6 (17.2)	90.7 (19.4)*	72.9 (13.3) / 73.1 (16.5)	81.1 (17.3) / 83.1 (19.1) ^{t}	81.1 (16.2) / 81.3 (17.4)	84.4 (16.1) / 83.3 (18.8)
Change Scores	-5.2 (14.8)	8 (17.2)**	-6.3 (11.8) / -4.4 (15.7)	-.02 (14.5) / 5.7 (19.2) ^{t}	-3.1 (17.2) / -0.6 (16.3)	3.4 (10) / -0.3 (20.4)
Vineland-II						
Subdomain(s)	Socialization		Communication		Daily Living Skills / Motor	
Entry Scores	76.4 (12.4)	77.5 (12.1)	73.8 (13.7)	79.4 (13.6)	83.5 (14.6) / 80.7 (11)	79.8 (14) / 78.8 (12.1)
Exit Scores	76.7 (12.7)	83.0 (9.7) ^{t}	79.3 (14.5)	87 (11.8)*	81.9 (15.4) / 76.9 (13.1)*	84.3 (13.4) / 84.3 (11.7)*
Change Scores	.3 (10.5)	5.5 (12.6)	5.5 (12.8)	7.6 (11.2)	-1.7 (17.1) / -3.8 (14.2)**	4.5 (12.7) / 5.5 (10.6)**

* *p* < .05, ** *p* < .01, *** *p* < .001, ^{t} *p* < .075.

Note: *n*'s are smaller than 70 because this analysis only compares the two groups that make up Summary Statement 1, those who change trajectories (progress categories c and d) versus those who begin below age-expected levels and do not change trajectories (progress categories a and b); it leaves out those who both begin and end at age-expected levels (progress category e)

Summary and Implications

The child assessments study provided an opportunity to consider how existing assessment tools are related to ratings derived from the COS process. While these kinds of analyses are quite common for validating assessment tools, the interpretation of findings from Study 2 are complicated by the measurement of different constructs across assessment tools. The BDI-2 and the Vineland-II measure developmental domains whereas the COS ratings reflect differences in the three functional child outcome areas. Therefore, strong relationships and high levels of correspondence in classification of children were not expected. However, we did expect to see some relationships between findings with these approaches.

The descriptive information about the COS from this study showed that the full range of ratings was used at entry and exit and that the percentages of children at the extremes was consistent with proportions anticipated. Some, but only a few children received ratings of 1 at entry on each of the outcomes and less than 15% of children entered with age-expected functioning on all three outcomes. These percentages were consistent with data from the BDI-2 and Vineland-II as well.

The child assessments study also examined the extent that COS ratings were associated with different child characteristics.   Children with higher scores on the ABILITIES Index had higher COS scores. Expected differences by primary disability also were found: ECSE children whose primary disability was speech or language impairment had higher COS ratings on average than children with other primary disabilities. Finally, children who entered at age expectations on all three outcomes both entered and exited with higher ratings on average on the COS than those who entered below age expectations on one or more outcomes.

Analysis of mean levels of assessment tool scores at each level of COS ratings produced data with expected patterns despite the small sample size. Likewise, detailed examination of correlations found that relationships between COS ratings and expected subdomains were around the levels that were predicted. In most cases, the correlations with expected subdomains were stronger than correlations that were not expected to be closely related to each outcome. However, differences in the relative strength comparing these correlations was weaker than found in many MTMM studies involving a single, unified construct and research on adult populations. Evidence also was found to support the idea that certain constructs (e.g., communication) were part of all three outcome areas.

Correspondence between outcome categories also was investigated with the progress categories produced by each of the three assessment approaches. Correspondence between the COS and the BDI-2 and Vineland-II tended to be fair or slight; however, correspondence between the BDI-2 and the Vineland-II was not substantial either, and only a little stronger than that observed with the COS. Follow-up analyses found that standardized scores on the BDI-2 and Vineland-II were not very effective at predicting the threshold for identifying children with age-expected functioning versus not. The COS process is designed for teams to build on and synthesize a wider array of information that includes, but goes beyond, assessment scores; this

complexity created challenges for empirical models to find clear linear relationships between assessment tools measuring domains and COS ratings of the three functional outcome areas. The reality that no specific, consistent decision-rules exist for converting the BDI-2 and Vineland-II to progress categories posed further challenges for interpreting low correspondence in progress category classification. While the small sample size available for these studies impacted methodological approaches that could be undertaken and influenced interpretation of the findings, it appears that several of these underlying methodological issues would continue to create challenges even with a much larger sample. Nevertheless, evidence showed that children whose ratings by COS teams indicated that the children changed their trajectories (i.e., the children represented in Summary Statement 1 percentages as showing greater than expected progress) did, on average, show positive changes between entry and exit on the BDI-2 and Vineland-II as well.

Taken together, despite the small sample sizes for the child assessments study and the methodological limitations in applying construct validity analytic approaches to the available data, the preponderance of evidence derived from the analyses conducted point in the direction of support for the COS process as a valid process to produce the intended information about children's functioning. In particular, variations in COS ratings and progress category classifications for children with independently obtained data about their skills and behaviors using standard assessment tools followed predictable patterns about their overall levels of functioning and the types of progress they make from entry to exit in EI and ECSE programs. As such, for accountability purposes, the COS process can yield credible data.

Section 5 — Study 3: Team Decision-Making Study – Design, Methods, Key Findings

Research Questions for Study 3

The team decision-making study was designed to examine implementation fidelity by coding videos of teams implementing the COS process. Study 3 addressed the following research questions:

1. What are the descriptive characteristics of ratings that emerge from the COS process?
2. What are the characteristics of teams and providers involved in the COS process?
3. What are the characteristics of team meetings to decide COS ratings?
4. What are the characteristics of the team process during COS meetings?
5. To what extent do teams implement the COS process with fidelity to guidance? This includes demonstrating understanding of the three outcomes, how to anchor skills to age-expected behavior, and understanding the COS rating criteria.
6. To what extent are ratings assigned that are consistent with the available evidence about the child in the video?
7. How does parent involvement on the team influence the COS process and ratings?
8. What other factors influence the COS process and ratings assigned?
9. To what extent is the evidence discussed during the COS process documented on the COS form?

Sample

State and Local Sample

For the team decision-making study, teams were instructed to video themselves conducting a COS discussion as they normally would. The researchers provided no additional training in the COS process to the members of the team. The project team received 138 videos, of which 131 were coded for analysis; 7 were excluded from coding because poor video or audio quality made them impossible to code. Of the initial 131 videos coded by the project team, 18 of the coded videos were excluded from analysis for the following reasons:

- On 4 videos, the coder identified that none of the time recorded on the video was focused on discussing a COS rating decision or talking about a child's functioning on the outcomes as necessary for the rating decision.

- On 10 videos, only one individual appeared. Although the person described what information he or she used, discussions he or she had with others that might contribute to the COS rating decision, and his/her thought process for selecting a rating, the videos yielded very little information about what occurred during that process.

- On 4 videos with more than one person appearing, the individuals were sharing and/or reviewing information individually rather than discussing the decision with each other at any point.

The final sample included in the analyses for the team decision-making study presented here was 113 videos. These videos were of COS team ratings from seven states that included 12 EI programs and 11 ECSE programs.[22]

Child Sample

The original study design called for children to be selected from the same local EI and ECSE program sample described in the General Methods section. The study plan had been to collect videotapes of 10 COS meetings, for a total of 360 team meeting videos (180 EI and 180 ECSE). Within each program, the 10 meetings were to be selected to represent a variety of child ages and types of disability specified by the project team. The 10 meetings within each program were to be 5 entry meetings (i.e., the first COS rating for a child) and 5 exit meetings (i.e., the final COSF rating for a child). Programs were encouraged to video-record COS entry meetings about children at a variety of age ranges.[23] For exit meetings, both EI and ECSE programs were encouraged to recruit those who had been in the program for varying lengths of time. The project expected

- about 40% of exit videos to include teams discussing children who participated in either EI or ECSE for less than 1 year and
- 60% of the videos to involve team discussions about children who had been in the program more than 1 year.

Both EI and ECSE were encouraged to recruit children reflecting a range of different types of disabilities as commonly observed in those programs nationally. In both entry and exit meetings, the following percentages of children were expected for EI:

- About 60% eligible for EI based on a developmental delay
- About 40% eligible for EI based on a diagnosed condition.

In ECSE, the following percentages of primary disability categories were expected across both entry and exit videos:

- About 40% with a primary disability of a speech or language impairment
- About 40% with a developmental delay that is not primarily a speech or language impairment
- About 20% with a primary disability category other than speech or language impairment or developmental delay.

---

[22] The original sampling plan called for 18 program sites, with six states each contributing data from three local areas.

[23] The original sampling plan for EI involved 40% of entry videos about children entering under 8 months of age, 20% between 8 and 20 months, and 40% entering after 20 months. This distribution was identified to approach the national distribution when children enter EI (Hebbeler et al., 2007). For ECSE, programs were encouraged to have at least 40% of entry videos about children under 48 months of age and 40% of entry videos about children older than 48 months, with flexibility in recruiting ages for the last 20% depending on the population served by that program.

These percentages were anticipated based on data described by Lazara, Danaher, Kraus, and Goode (2009) indicating the frequency of speech or language impairment and developmental delay as primary disability categories in the preschool special education population.

The actual sample sizes were much smaller, and the characteristics of the children are shown in Exhibit 5.1. Of the 113 videos,

- 65% of the videos were of meetings to decide entry ratings, and 35% were meetings for exit ratings.
- 56% of the videos were of team meetings about children in EI, and 44 % were about children in ECSE.
- Few statistically significant differences were observed between entry and exit meetings or between EI and ECSE meetings across a range of demographic and descriptive characteristics.
- As expected, children were older at the time of exit meetings than entry meetings, and children in ECSE were older than children in EI.
- For EI entry meetings, 37% of children discussed at the meeting were younger than 1 year, 37% were 1 to 2 years old, and 26% were 2 to 3 years old.
- For ECSE entry meetings, 29% of children discussed at the meeting were not quite 3 years old, 51% were 3 to 4 years old, and 20% were 4 years old or older.
- Both EI and ECSE had fairly close to expected percentages of children at exit who had been in the program less than 1 year, with 40% among EI exit videos and 33% among ECSE exit videos.
- Both EI and ECSE had somewhat higher percentages of children in the developmental delay category than expected (EI 73%, ECSE 52%), but for both programs there was still a good spread across expected eligibility or primary disability categories that is fairly similar to expected percentages.
- No statistically significant differences in ABILITIES Index scores were observed between entry and exit meetings or between EI and ECSE meetings.

Exhibit 5.1   Child Sample for Team Decision-Making Study

Characteristics	Meeting Timing		Program		Overall n (%)
	Entry Videos n (%)	Exit Videos n (%)	EI n (%)	ECSE n (%)	
Number of children	73 (65)	40 (35)	63 (56)	50 (44)	113 (100)
Program					
EI	38 (52)	25 (63)			63 (56)
ECSE	35 (48)	15 (38)			50 (44)
Gender					
Female	22 (30)	14 (35)	22 (35)	14 (28)	36 (32)
Male	51 (70)	26 (65)	41 (65)	36 (72)	77 (68)
Race/ethnicity†					
African American/Black	15 (21)	3 (8)*	10 (16)	8 (16)	18 (16)
Hispanic/Latino	12 (16)	15 (38)	17 (27)	10 (20)	27 (24)
Caucasian/white	38 (52)	18 (45)	29 (46)	27 (54)	56 (50)
Other	8 (11)	4 (10)	7 (11)	5 (10)	12 (11)
Age at COS meeting					
Mean (SD) in months	28.7 (15.9)	46 (15.9)***	23.5 (12.3)	49.0 (13.0)**	36.3 (12.7)
< 6 months	10 (14)	0 (0)	10 (16)		10 (9)
6-11 months	4 (5)	0 (0)	4 (6)		4 (4)
1 year	14 (19)	1 (3)	15 (24)		15 (13)
2 years	20 (27)	15 (38)	25 (40)	10 (20)^	35 (31)
3 years	18 (25)	10 (25)	9 (14)	19 (38)	28 (25)
4 years	5 (7)	2 (5)		7 (14)	7 (6)
5 years or older	2 (3)	12 (30)		14 (28)	14 (12)
Disability type in EI					
Diagnosed condition	12 (16)	5 (13)	17 (27)		17 (15)
Developmental delay	26 (36)	20 (50)	46 (73)		46 (41)
Disability type in ECSE					
Developmental delay	16 (22)	10 (25)		26 (52)	26 (23)
Speech-language impaired	12 (16)	3 (8)		15 (30)	15 (13)
Other	7 (10)	2 (5)		9 (18)	9 (8)
ABILITIES Index total score (reversed)					
Mean (SD)	101.6 (12.2)	102.2 (10.4)	101.0 (13.2)	102.7 (9.0)	101.9 (11.1)
Low (< 100)	23 (31)	11 (28)	21 (33)	13 (26)	34 (30)
Moderate (100-107)	22 (30)	18 (45)	19 (30)	21 (42)	40 (35)
High (108-114)	28 (38)	11 (28)	23 (36)	16 (32)	39 (35)

Exhibit 5.1   Child Sample for Team Decision-Making Study (concluded)

Characteristics	Meeting Timing		Program		Overall n (%)
	Entry Videos n (%)	Exit Videos n (%)	EI n (%)	ECSE n (%)	
Length of time In program *(Between COS entry and COS exit— for exit videos only)*					
Mean (*SD*) in months		18.6 (10.6)	16.6 (9.7)	21.9 (11.4)	18.6 (10.6)
Range in months		5-36	5-35	8-36	5-36
< 1 year		15 (38)	10 (40)	5 (33)	15 (38)
1 to 2 years		12 (30)	9 (36)	3 (20)	12 (30)
2 years or more		13 (33)	6 (24)	7 (47)	13 (33)

* *p* < .05, ** *p* < .01, *** *p* < .001 for comparisons for EI and ECSE and entry and exit meetings. All comparisons, including those between entry and exit meetings, represent different groups of children.

† Overall chi square is significant, crosses all four categories of race/ethnicity (χ^2 = 7.8411, *p* = .0494).

Note: In several ECSE programs, entry COS meetings were held as part of meetings that occurred before the child's third birthday, thus allowing the child to begin receiving services as soon as he/she turned 3 years of age.

Methods for Study 3

Measures

Three forms were developed by the project team for use in the team decision-making study: Child and Meeting Information Form, Provider Information Form, and Videotape Coding Protocol. Copies of these three forms are in Appendix L.

- The one-page **Child and Meeting Information Form** was used to collect information about the child being discussed and the meeting itself. Information about the child was the date of birth, gender, ethnicity, primary language, other languages spoken, and reason for eligibility (for EI) or primary disability (for ECSE). Information about the meeting was the names and roles (e.g., speech therapist) of those who contributed information and who was present, as well as the date of the meeting. This form was completed by the local program staff.

- The **Provider Information Form** was collected only once for each provider, regardless of how many times the provider appeared in different videos. Provider information consisted of years of experience providing EI services, years of experience with the COS process, amount of training on the COS process, and self-reported understanding about several key COS concepts.

- The video of the meeting was coded by the research team using the **Videotape Coding Protocol**. Along with each video, programs were asked to submit a signed statement of informed consent from the parent and all present at the meeting, a child and meeting information form, a provider information form for each person in the video, a copy of the program's COS form where the team recorded its decision, and a completed ABILITIES Index[24] about the child.

Data Collection and Coding Procedures

SRI provided local programs with a video camera, tripod, SD cards, and instructions for filming. Programs were sent information packets and invited to attend informational webinars held at various times. Project staff explained how to select team meetings for the study, shared information about the forms to complete, and described how to operate the video camera. In addition, information was given about how to seek parent and provider consent for the videos, along with information for parents about the study and the necessary consent forms. Programs sent the recorded video, a copy of the completed COS form, the consent forms, the Provider Information Forms, and the Child and Meeting Information Form for each focal child to the research team.

---

[24] The same measure and subareas were used in the team decision-making study as in the child assessments study. Internal consistency of this scale was examined for the team decision-making study sample ($n = 113$) and alphas of .90, .89, .86, and .92 were found respectively for the total scale score, the Social/Communication subarea, the Cognitive/Communication subarea, and the Structural Integrity subarea.

A coding system that the ECO Center had used for an examination of videos of 55 COS team meetings was further refined for this study (Gould & Walker, 2008)). The coding system captured information about how teams conducted and documented the meeting based on observations from the video and review of the completed COS form. It focused on the validity claims (knowledge of the scoring criteria, the extent that discussion reflected understanding of the outcomes and age-appropriate child development, etc.), with most indicators coded on a 3-point scale of whether the indicator was *completely present*, *partially present*, or *not at all present* in the video. The number and professional roles of team members along with the presence of a parent were coded. Two research analysts jointly coded the first 20 videos to ensure consistent application of the coding categories and refine the codebook. Numerous iterations of the video coding protocol evolved as videos were reviewed and coded by the research analysts and the additional information necessary to address the validity questions was identified and required coding of additional data. The final Videotape Coding Protocol captured over 150 data points. A summary of the content of the coding protocol is presented in Exhibit 5.2, and a copy of the protocol is in Appendix L.

After the coding protocol and codebook were refined and coders were consistently identifying similar ratings during joint coding, the next three videos were independently coded by the two research analysts. After each of these videos was coded, the results were compared and discussed to reach consensus by the coders. All previously coded videos then were reviewed and coded again to ensure consistent application of the final codebook approach. Approximately every twentieth video thereafter was independently coded by both research analysts to detect any drift in coding. Disagreements were subsequently discussed and resolved through consensus to produce one set of codes for the final analysis. Coders had over 93% agreement across all codes on every video that was double-coded. These instances of independent double ratings also were used to obtain estimates of interrater reliability.

Exhibit 5.2   Summary of Content Examined in Videotape Coding Form

Meeting and Participant Information: *Description of type of meeting where COS was decided, participants involved, and their backgrounds.*

- Type of meeting where ratings were decided (e.g., embedded in IFSP/IEP, COS-only meeting)
- Number and roles of individuals who contributed information to the rating and/or participated in the rating decision
- Years of experience providing services to children with disabilities and to children without disabilities
- Entry or exit meeting
- Number of parents/guardians present at meeting
- Hours of COS training received by providers
- Length of meeting and length of time spent discussing each outcome
- Number of COS meetings the providers have participated in
- Providers self-reported knowledge about the COS process

Informing and Engaging Family Members: *Information explained to families and their participation in COS process.*

- Providers explain to families why outcomes data are collected.
- Providers describe to families the meaning of the three child outcomes.
- Families share information about their child's functioning without prompting from providers.
- Providers review with families the skills expected at the child's age and the sequence in which those skills develop.
- Providers invite family members to share observations and input during the meeting.

Team Process and Consensus: *General interaction approach and dialogue between members at the COS meeting.*

- Extent of relevant contributions and dialogue between team members
- Amount of family input described or shared (minimal, considerable)
- Format of input received from family members (e.g., professional shared based on separate discussions, family participation in team meeting)
- Consensus agreement on the COS ratings and the rationale for the ratings
- Overall rating of team process

COS-Specific Process: *Team incorporates best practices and understanding of COS-related content in meeting.*

- Team discussed the child's functional use of skills.
- Team described skills the child has not yet mastered.
- Team considered more than one rating.
- Team considered the child's functioning across multiple settings and situations.
- Team correctly age-anchored specific skills.
- Team explicitly stated a rationale for rating for reach outcome.
- Team discussed skills relevant to each of the three outcomes.
- Team referenced one or more specific assessment tools.
- Team showed no explicit intent to alter ratings.
- Team considered the full breadth of skills associated with the outcome.
- Team referenced or used the decision tree in the process of deciding the rating.
- Overall rating of COS-specific process
- Team considered the child's skills with the appropriate depth.
- No misunderstanding of rating criteria was evident.

Ratings: *Team ratings are consistent with specific rating or range of ratings determined by coding after reviewing the video and COS form.*

Documentation: *Review of COS form to consider if paperwork is complete and provides evidence that justifies the rating.*

- COS form is complete, with evidence documented for each outcome.
- COS evidence listed is identified with the appropriate outcome area.
- Evidence documented is consistent with and/or sufficient to justify the rating.

Data Analysis for Study 3

Descriptive analyses were conducted initially for each research question. Depending on the nature of the observations (categorical, ordinal, or interval), the analyses included univariate statistics—frequency distributions, traditional statistical measures (mean, median, percentiles, standard deviation, skewness, and kurtosis), and box plots, as well as bivariate statistics—cross-tabulations, scatterplots, measures of association such as correlations (Pearson, Spearman, etc.), and chi-square statistics.

When the coding variable was dichotomous, we determined characteristics that were associated with variation in the probability of a positive value. We used mixed-effects logistic models where the covariates, for example, included the characteristics of the child under consideration. We first tested for the joint statistical significance of groups of covariates within a domain (e.g., the joint contribution of measures of characteristics of the child or measures of the structure and length of the discussion). If the covariates were jointly significant, we conducted further regressions for individual covariates within that set to determine which covariates appeared to be most influential. We repeated many of these same kinds of analyses to see whether various predictor variable were influential in predicting team ratings for the child above and beyond what would be predicted by the ABILITIES Index information about the extent the child's disability was influencing his/her functioning. Either polytomous categorical or ordinal logistic models were used with categorical data.

In addition, we examined interaction effects (both within and across domains) when we had reason to believe (on the basis of prior work or the literature) that the interactions may be influential. Because the purpose of these regression analyses was to identify areas where it would be desirable to refine and improve the COS form and the supporting materials, and because the consequences of a false positive (i.e., consideration of improvements to the tool or training materials) were outweighed by the consequences of a false negative (failing to identify and correct a deficiency), we did adjust for multiple comparisons.

Space constraints limited the number of specific analyses and findings described in this report. We concentrated on providing information so that others reviewing COS information and conducting ongoing validity analyses could compare their findings with these. For more information about the specific kinds of analyses undertaken with certain kinds of data or in response to specific research questions, please contact the report authors.

Key Results for Study 3: Team Decision-Making Study

Key results from the team decision-making study are organized by research question. Descriptive findings about the ratings derived from coding the videos are presented first. Next, more information is provided about the teams, providers, and the team process observed in the videos. Then findings about the fidelity to COS guidance and the quality of COS-specific practices observed in videos are presented. These include data on the consistency between team ratings and ratings provided by external coders based on information evident in the video. The final sections examine the influence that a variety of demographic, team, and other factors have on the COS ratings and the COS process and describes how teams documented decisions from the COS process on forms.

Distributions of COS Ratings in Videos

Described here are the distributions of COS ratings observed in coded videos. We present mean ratings, the range of ratings observed, the percentage of COS ratings that identify children at age-expected levels, and correlations between the outcomes using COS ratings at entry and exit.

Mean COS Ratings

The mean COS ratings for videos in the team decision-making study (Exhibit 5.3) show that

- Mean ratings in exit meetings were consistently higher than mean ratings at entry meetings.
- These differences reached statistical significance for all but positive social relationships in EI.
- These differences are consistent with other data from ENHANCE showing higher exit COS scores on average at program exit than at entry to EI or ECSE services.

Exhibit 5.3 **Mean COS ratings in Team Decision-Making Study Overall and for EI and ECSE in Entry and Exit Meetings**

	Entry Meetings (*n* = 73)			Exit Meetings (*n* = 40)		
	Positive Social Relationships	Knowledge and Skills	Action to Meet Needs	Positive Social Relationships	Knowledge and Skills	Action to Meet Needs
	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
Overall (entry *n* = 73, exit *n* = 40)	4.8 (1.7)	4.4 (1.6)	4.5 (1.4)	5.5 (1.3)	5.2 (1.2)	5.5 (1.4)
EI (entry *n* = 38, exit *n* = 25)	4.7 (2.0)	4.1 (1.7)	4.3 (1.5)	5.3 (1.3)	5.0* (1.3)	5.2* (1.4)
ECSE (entry *n* = 35, exit *n* = 15)	4.9 (1.3)	4.7 (1.5)	4.7 (1.3)	5.8* (1.4)	5.5* (1.1)	6.0** (1.4)

* $p < .05$, ** $p < .01$, *** $p < .001$ for comparing entry and exit ratings for each outcome within overall, EI, or ECSE respectively. Note that entry and exit meetings represent different groups of children.

Range of COS Ratings

The 113 videos in the study spanned the full range of COS ratings on all the outcomes (Exhibit 5.4). The full range of 1–7 ratings was observed on each of the three outcomes among the 73 videos of entry meetings. Among entry meetings, 4%, 6%, and 3% of videos showed team ratings of 1 on positive social relationships, knowledge and skills, and action to meet needs, respectively.

Exhibit 5.4   Percentage of Children with Various COS Ratings in Overall Sample for Team Decision-Making Study (*n* = 113)
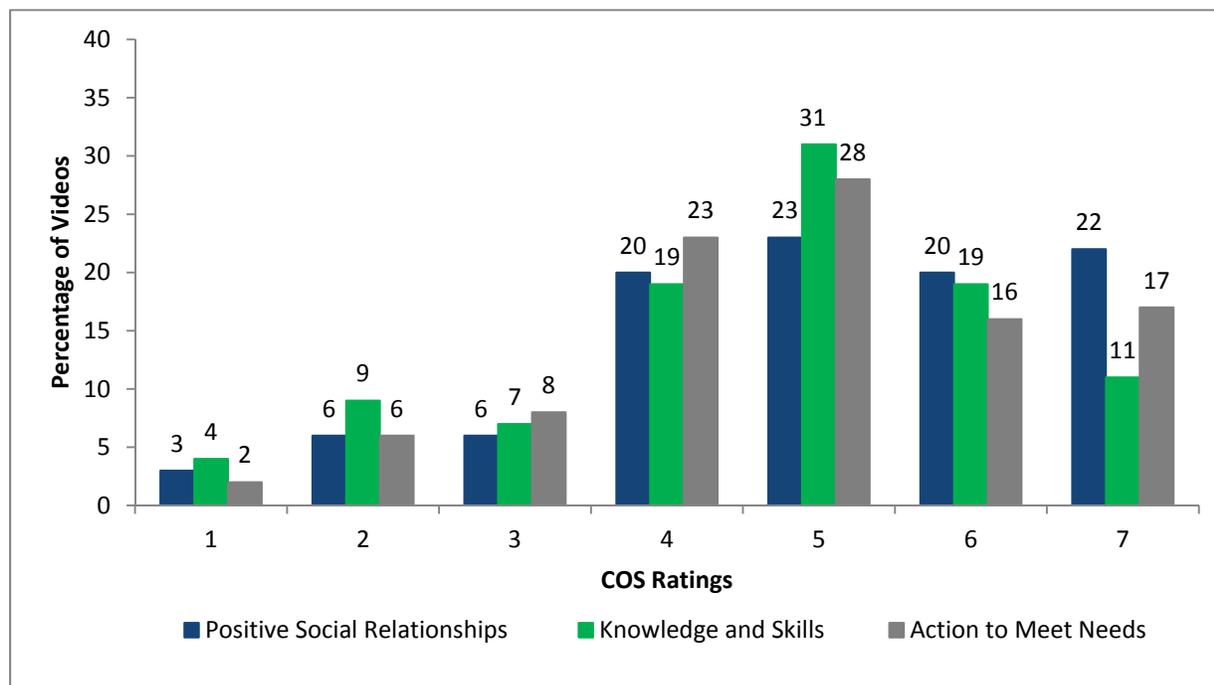
Percentage of COS Ratings at Age-Expected Levels

Of the 73 videos of meetings at program entry, 14% of the children (10) entered at age-expected levels on all three outcomes (i.e., 14% had ratings of 6 or 7 on all three outcomes at the entry meeting).

COS Ratings: Correlations Between Outcomes

The videos for any one child provided ratings at only one point in time. Correlations were examined between the ratings on the three outcomes in videos of entry meetings (Exhibit 5.5) and videos of exit meetings (Exhibit 5.6). In both cases, correlations between outcomes were strong.

Exhibit 5.5   COS Rating Correlations Between Outcomes in Entry Videos from Team Decision-Making Study (*n* = 73)

Team Decision-Making Study Entry Videos – COS Ratings	Positive Social Relationships	Knowledge and Skills	Action to Meet Needs
Positive social relationships			
Knowledge and skills	.75		
Action to meet needs	.78	.81	

Exhibit 5.6   COS Rating Correlations Between Outcomes in Exit Videos from Team Decision-Making Study (*n* = 40)

Team Decision-Making Study Exit Videos – COS Ratings	Positive Social Relationships	Knowledge and Skills	Action to Meet Needs
Positive social relationships			
Knowledge and skills	.80		
Action to meet needs	.80	.75	

Characteristics of Teams and Providers Involved in COS Process

Description of Teams Involved in COS Process

No sampling requirements were included for specific team features in the team decision-making study. Programs were instructed to use "whatever team would normally meet to identify the child's COS ratings" for children and families with signed consent for this study. Exhibit 5.7 shows the characteristics of teams who were involved in the COS ratings decisions for this study.

- Most teams included three or four people in the COS rating discussion, counting both professionals and family members who were present.
- About half of the teams had a family member present for the COS rating discussion and most had a service coordinator present.
- Often, documentation indicated that one to two additional people contributed information to the rating decision, but were not present for the COS rating discussion.
- On teams were family members were not present for the rating discussion, most contributed information to inform the rating.
- Given that many of the programs were fairly small, 76% of the teams had one or more providers who also had appeared in one or more other videos.
- About half of teams had at least one team member who had 9 or more hours of COS training, but for about 1/3 of teams the average amount of COS training received by professionals on the team was less than three hours. This was more likely to occur at exit than entry meetings.
- About half of the meetings were held separately, specifically for the COS rating decision. About 20% held COS rating discussions at the end of the IFSP or IEP meeting and a few more embedded the discussion within the IFSP/IEP meeting.

Exhibit 5.7. Team Sample for Team Decision-Making Study

	Timing of COS Meeting		Program		Overall/ All Meetings Combined *n* (%)
	Entry Meetings *n* (%)	Exit Meetings *n* (%)	EI All Meetings *n* (%)	ECSE All Meetings *n* (%)	
Number of children	73 (65%)	40 (35%)	63 (53%)	50 (44%)	113 (100%)
Number of states	7	7	6	4	7
Number of programs	21	21	12	11	21
Family Member Present for COS Discussion	38 (52%)	21 (53%)	32 (51%)	27 (54%)	59 (52%)
Service Coordinator Present for COS Discussion	61 (84%)	37 (93%)	56 (89%)	42 (84%)	98 (87%)
Total Number Present for COS Discussion (Family Members and Professionals) Mean (SD)	4 (2.2)	3 (1.5)	3 (1.2)	4 (2.6)	4 (1.9)
Total Number of Professionals Present for COS Discussion Mean (SD)	3.3 (1.9)	2.5 (1.3)	2.5 (0.7)	3.6 (2.3)	3 (1.7)
1	5 (7%)	4 (10%)	1 (2%)	8 (16%)	9 (8%)
2	23 (32%)	23 (58%)	36 (57%)	10 (20%)	46 (41%)
3	27 (37%)	8 (20%)	21 (33%)	14 (28%)	35 (31%)
4	7 (10%)	1 (3%)	4 (6%)	4 (8%)	8 (7%)
5 or more	11 (15%)	4 (10%)	1 (2%)	14 (28%)	15 (13%)
Family Member contributed information* to the COS Decision	72 (99%)	39 (98%)	61 (97%)	50 (100%)	111 (98%)
Total Number who contributed information* to the COS Decision (Family Members and Professionals) Mean (SD)	5 (2.2)	4 (1.6)	4 (1.0)	6 (2.4)	5 (1.7)

Exhibit 5.7. Team Sample for Team Decision-Making Study (concluded)

	Timing of COS Meeting		Program		Overall/
	Entry Meetings n (%)	Exit Meetings n (%)	EI All Meetings n (%)	ECSE All Meetings n (%)	All Meetings Combined n (%)
Meeting Format					
In IFSP/IEP	5 (7%)	2 (5%)	6 (10%)	1 (2%)	7 (6%)
At end of IFSP/IEP	18 (25%)	6 (15%)	12 (19%)	12 (24%)	24 (21%)
Separate COS meeting	38 (52%)	32 (80%)	36 (57%)	34 (68%)	70 (62%)
At end of another/ assessment meeting	10 (14%)	0 (0%)	7 (11%)	3 (6%)	10 (9%)
Other	2 (3%)	0 (0%)	2 (3%)	0 (0%)	2 (2%)
Amount of training for person on the COS decision team with the most training					
< 3hrs	6 (8%)	10 (25%)	13 (21%)	3 (6%)	16 (14%)
3-4hrs	12 (16%)	12 (30%)	13 (21%)	11 (22%)	24 (21%)
5-8hrs	17 (23%)	3 (8%)	10 (16%)	10 (20%)	20 (18%)
≥ 9hrs	38 (52%)	15 (38%)	27 (43%)	26 (52%)	53 (47%)
Teams with at least one person with 31 or more COSFs and has at least 9 hours of training n (%)	32 (44%)	14 (35%)	25 (40%)	21 (42%)	46 (41%)
Teams where average amount of COS-related training received by members on the team was less than 3 hours n (%)	16 (22%)	22 (55%)	21 (33%)	17 (34%)	38 (34%)
Professional roles of those present for COS rating discussion					
Early Interventionist/Teacher	62 (85%)	32 (80%)	53 (84%)	41 (82%)	94 (83%)
Therapist	63 (86%)	27 (68%)	51 (81%)	39 (78%)	90 (80%)
Service Coordinator/Case Manager	31 (42%)	17 (43%)	32 (51%)	16 (32%)	48 (43%)
Educational Psychologist/Social Worker	24 (33%)	7 (18%)	8 (13%)	23 (46%)	31 (27%)
Other	22 (30%)	8 (20%)	9 (14%)	21 (42%)	30 (27%)

* Some teams had individuals contribute information to the rating decision who were not present at the meeting. These numbers describe the characteristics of the full team who provided input at any time toward deciding the COS rating.

Characteristics of Providers involved in COS Decision Teams

Providers involved in the COS rating decisions for this study provided information about their backgrounds (Exhibit 5.8) and perceived understanding of and confidence with the COS process (Exhibit 5.9).

- 208 providers were involved across the 113 videos.

- A range of different professionals participated. Therapists were more likely to be present at EI meetings than ECSE meetings $(p < .0001)$ and psychologists or social workers were more likely to be present at exit meetings than entry meetings $(p < .0001)$.

- Most providers had quite a bit of experience working in EI or ECSE, with about half working with young children with disabilities 11 years or more and only 6% who had been working in the field less than a year.

- About 80% of providers indicated that they also had worked with young children without identified special needs which may be important for their perspectives on anchoring what behavior is age-expected.

- 42% had participated in more than 50 COS process rating decisions. EI providers $(p < .01)$ and providers at entry meetings $(p < . 05)$ were more likely to indicate participation in more past COS meetings.

- Eight percent of the providers in the videos had not previously participated in the COS process, or at least didn't recognize it by the name when queried. Lack of prior experience was more common among providers at exit meetings and slightly more frequent among ECSE providers.

- Only 13% had both considerable experience with the COS process and 9 or more hours of training about the COS.

- No statistically significant differences were found in self-rated understanding among EI versus ECSE providers or providers at entry versus exit meetings for core skills needed for the COS process (items shown in Exhibit 5.9).

We had intended that the providers and teams involved in ENHANCE would be from programs at the highest level of implementation. That is, providers were expected to be highly trained and have experience implementing the COS process. The data in Exhibits 5.8 and 5.9 indicate that the sample of providers included considerable variation in provider training, experience, and perceived self-understanding of key aspects of the COS process. Variation in provider familiarity with the COS could be expected to impact the characteristics of the meetings that were videoed.

Exhibit 5.8   Provider Sample for Team Decision-Making Study

| Characteristics | Timing of COS Meeting | | Program | | Overall/ All Meetings Combined (*n* = 208) *n* (%) |
	Entry Meetings (*n* = 90) *n* (%)	Exit Meetings (*n* = 118) *n* (%)	EI All Meetings (*n* = 111) *n* (%)	ECSE All Meetings (*n* = 97) *n* (%)	
Professional role					
Early Interventionist or Teacher	58 (37%)	32 (63%)***	48 (43%)	42 (43%)	90 (43%)
Service/Family Resource Coordinator	60 (38%)	16 (31%)	45 (41%)	31 (32%)	76 (37%)
Therapist	25 (16%)	9 (18%)	32 (29%)	2 (2%)****	34 (16%)
Psychologist/Social Worker	16 (10%)	0 (0%)*	1 (1%)	15 (15%)****	16 (8%)
Other	14 (9%)	2 (4%)	7 (6%)	9 (9%)	16 (8%)
Length of time providing services to young children with disabilities					
Less than 1 year	8 (5%)	4 (8%)	5 (5%)	7 (7%)	12 (6%)
1-2 Years	12 (8%)	7 (14%)	14 (12%)	5 (5%)	19 (9%)
3-5 Years	19 (12%)	12 (24%)	20 (18%)	11 (11%)	31 (15%)
6-10 Years	36 (23%)	10 (20%)	20 (18%)	26 (27%)	46 (22%)
11 Years or More	82 (52%)	18 (35%)	52 (47%)	48 (49%)	100 (48%)
Worked with young children without disabilities - Yes *n* (%)	124 (80%)	42 (84%)	75 (80%)	91 (82%)	166 (81%)
Number of COSF's participated in					
Zero	9 (6%)	7 (14%)*	4 (4%)	12 (12%)**	16 (8%)
1-10	19 (12%)	7 (14%)*	11 (10%)	15 (15%)**	26 (13%)
11-30	30 (19%)	15 (29%)*	19 (17%)	26 (27%)**	45 (22%)
31-50	24 (15%)	9 (18%)*	22 (20%)	11 (11%)**	33 (16%)
More than 50	75 (48%)	13 (25%)*	55 (50%)	33 (34%)**	88 (42%)

Exhibit 5.8   Provider Sample for Team Decision-Making Study (concluded)

Characteristics	Timing of COS Meeting		Program		Overall/ All Meetings Combined (*n* = 208) *n* (%)
	Entry Meetings (*n* = 90) *n* (%)	Exit Meetings (*n* = 118) *n* (%)	EI All Meetings (*n* = 111) *n* (%)	ECSE All Meetings (*n* = 97) *n* (%)	
Number of hours of training received					
None	8 (5%)	9 (18%)*	7 (6%)	10 (10%)*	17 (8%)
Less than 1 hour	15 (10%)	10 (20%)*	13 (12%)	12 (13%)*	25 (12%)
1-2 hours	40 (26%)	13 (25%)*	31 (28%)	22 (23%)*	53 (26%)
3-4 hours	34 (22%)	10 (20%)*	22 (20%)	22 (23%)*	44 (21%)
5-8 hours	30 (19%)	4 (8%)*	18 (16%)	16 (17%)*	34 (16%)
9-15 hours	11 (7%)	2 (4%)*	9 (8%)	4 (4%)*	13 (6%)
More than 15 hours	18 (12%)	3 (6%)*	11 (10%)	10 (10%)*	21 (10%)
Number and percent of providers with 31 or more COSFs and had 9 hours or more of training	23 (15%)	4 (8%)	17 (15%)	10 (10%)	27 (13%)

* *p* < .05, ** *p* < .01, *** *p* < .001, **** *p* < .0001.

Exhibit 5.9   Providers' Self-Reported Understanding of COS Background Information in Team Decision-Making Study

Category	Not well/Not at all well	Well	Very well
Understand skills in each of the three outcome areas	4	58	38
Understand the 7 COS rating points	6	60	33
Understand the degree skills and behaviors are age appropriate	3	56	41
Report understanding very well for all three questions			29

Percentage of Providers

■ Not well/Not at all well   ■ Well   ■ Very well
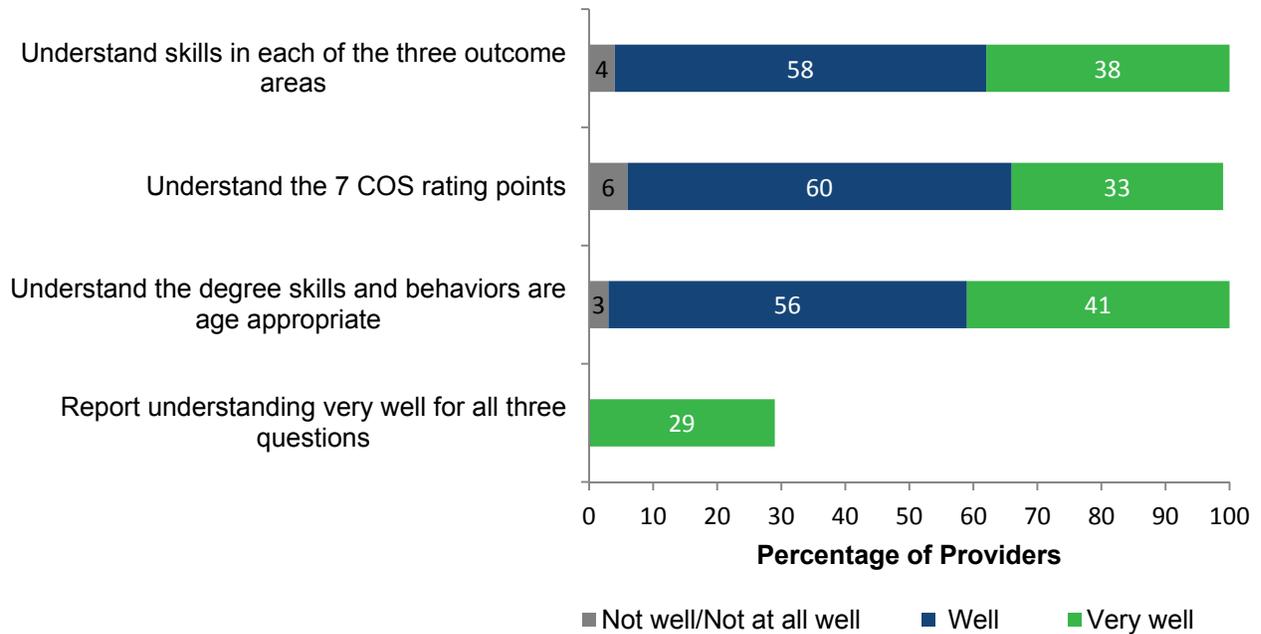
Characteristics of Team Meetings to Decide COS Ratings

COS Meeting Format

COS team decisions were made in a variety of meeting formats:

- Most of the videos (62%) were of stand-alone meetings held specifically for the COS process.

- One-fourth of the videos (21%) showed the COS process at the end of an IFSP or IEP meeting. In that case, discussion about the child's present levels of development or functioning also may have occurred earlier in the meeting with the same participants, but it was not included on the video.

- Fewer videos showed the COS process occurring at the end of an evaluation/assessment team meeting (9%).

- The least frequent meeting format among the videos was COS decision-making embedded within the larger IFSP or IEP team meeting (6%). In these videos, the entire IFSP or IEP meeting was video-recorded.

No statistically significant differences were found between EI and ECSE teams in the format of the meeting in which the COS process was embedded. However, the format of the meeting for the COS was related to whether or not a parent/family member was included on the team making the COS rating decision.

- Almost two-thirds of COS-only (67%) and evaluation/assessment (60%) team meetings did *not* include a parent or family member.
- In contrast, all the video embedded in or at the end of the IFSP/IEP meetings included a parent or family member.

Length of COS Process Discussion

The COS process varied depending on the participants involved as well as their experiences working together in prior COS process meetings, but some common features of these meetings are the following:

- The COS process includes some introduction to a need for identifying a COS rating and sometimes why the information is collected and/or what happens with the information.
- There also is an introduction to the content area of the first outcome. For example, the facilitator might ground the group in what is meant by the outcome area of positive social relationships.
- Next, the team discusses the child's level of functioning with regard to different skills and behaviors related to the outcome area. The team considers the child's functioning relative to what is expected for the child's age.
- Then the child's functioning is considered relative to the rating criteria for 1–7. The team discusses possible ratings, often sharing a rationale for a potential rating, and the group reaches consensus about a rating decision on that outcome.
- This process is then repeated for outcomes 2 (acquiring and using knowledge and skills) and 3 (taking appropriate action to meet needs).

The coder recorded the amount of time the team spent on the COS process. Time spent on each outcome decision was recorded to the nearest minute, with rounding. Overall key COS process time was based on the sum of rounded time for each outcome, plus any time the facilitator spent introducing the activity or wrapping up the meeting. Time was deducted if team discussion concerned a tangent unrelated to the child's functioning, ratings, or the COS process. When the COS process was embedded in longer meetings (e.g., the IFSP or IEP), the coder summed time spent on the key elements of the process, including discussions about the child's functioning. Other topics, such as planning for services and discussing schedules, were not included in the time recorded.

Results (see Exhibit 5.10) were the following:

- The COS process ranged in length from 1 to 37 minutes.
- On average, it was 10 minutes ($SD = 7.5$ minutes).
- No statistically significant difference was found in key COS meeting length between EI and ECSE programs (EI meetings, 11 minutes [$SD = 7.9$]; ECSE meetings, 9 minutes [$SD = 6.8$]).
- About two-fifths of the videos of the COS process were 5 to 9 minutes long.

The findings about the length of time for the key COS process in the team decision-making study complement the reported amount of time spent by those who completed the provider

survey (Study 1). On the provider survey, 36% reported 1–15 minutes, 35% reported 16–30 minutes, and 18% reported 31–45 minutes to complete the COS discussion and document the decision with any necessary evidence on the COS form. No information is available from either study about the amount of time required to document the COS rating decision. However, it seems that by either measure, teams were keeping conversations and documentation fairly brief.

Exhibit 5.10   Length of Key COS Process Across Outcomes in Team Decision-Making Videos *(n* = 113)
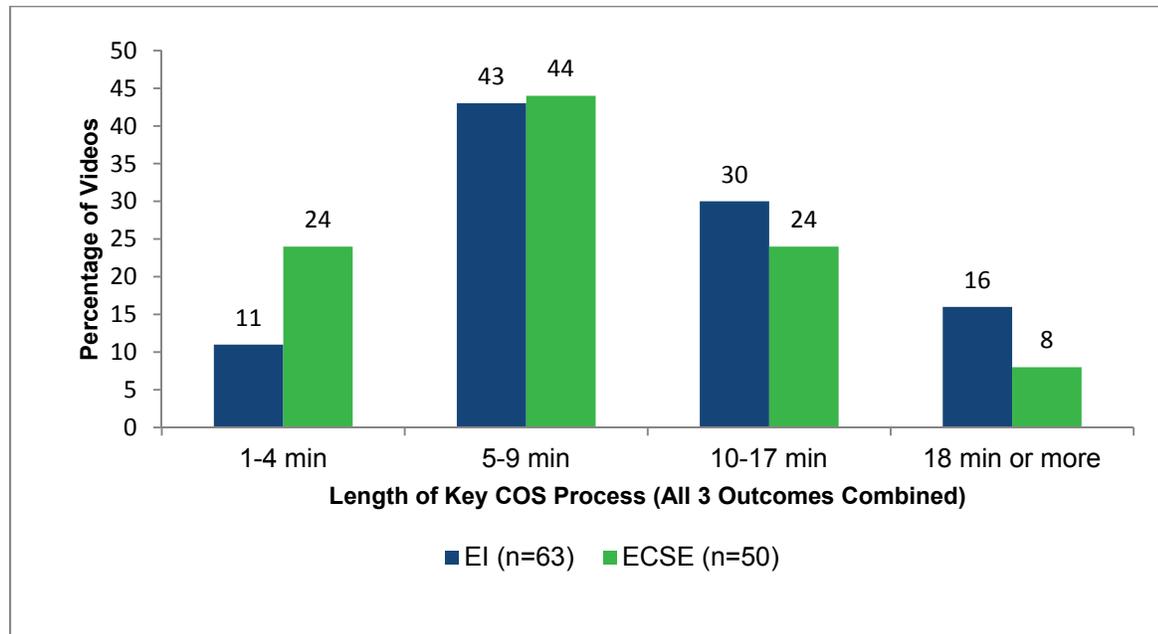
Team Process during COS Meetings

During video coding, the coder rated the overall quality of the team process on a 5-point Likert-type scale, with 1 being the *lowest quality* and 5 being the *highest quality*. The rating was made after considering a series of items about the quality of interactions between team members in the video, including the extent to which the team had rich dialogue and considered input from all team members. Exhibit 5.11 shows the distribution of ratings of the team process in videos.

- A majority of overall quality ratings were in the mid to high range.
- The average overall quality of the team process was 3.6 (*SD* = 1.2).
- No statistically significant differences were found in overall quality of the team process between EI and ECSE or between entry and exit meetings.

Exhibit 5.11   Overall Quality of Team Process During Meetings (*n* = 113)
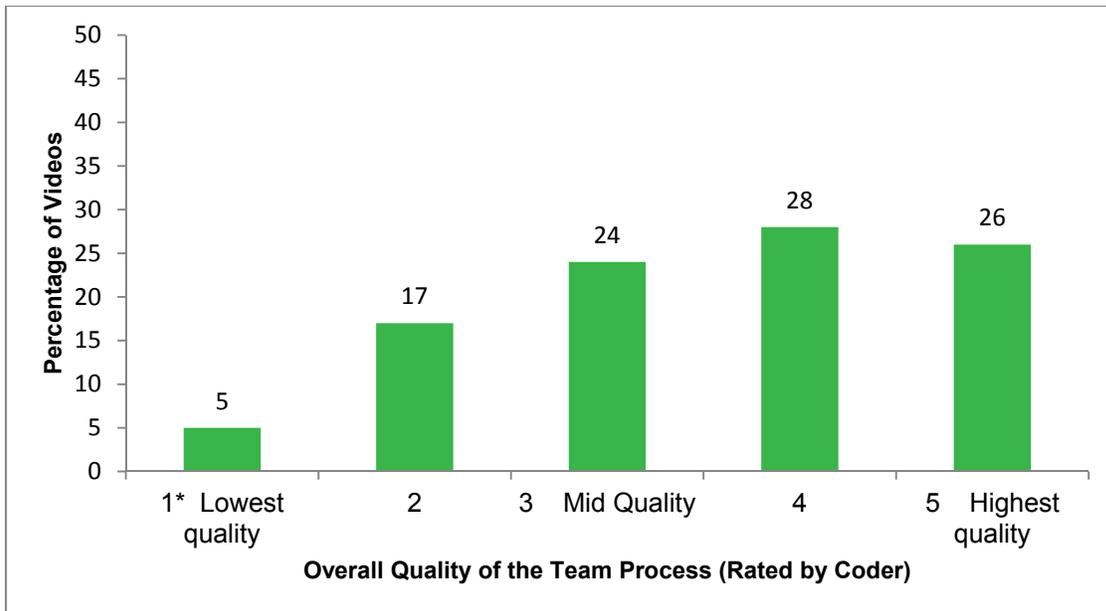
*   Two videos were coded as "Can't tell" for the quality of their overall team process. Given that the team process was so minimal or poor that it was not visible enough to code, these two videos were clustered with the lowest quality group.

The overall quality of the team process was related strongly to the length of time teams spent on the key COS process (χ^2 = 65.9, *p* < .0001; *r* = .57, *p* < .0001; Exhibit 5.12 shows the relationships between these two variables). Teams that spent longer on the COS process were rated by the coder as having higher levels of quality in their overall team process.

Exhibit 5.12   Relationship Between Length of Time of COS Process and Overall Quality Ratings on Videos (*n* = 113)
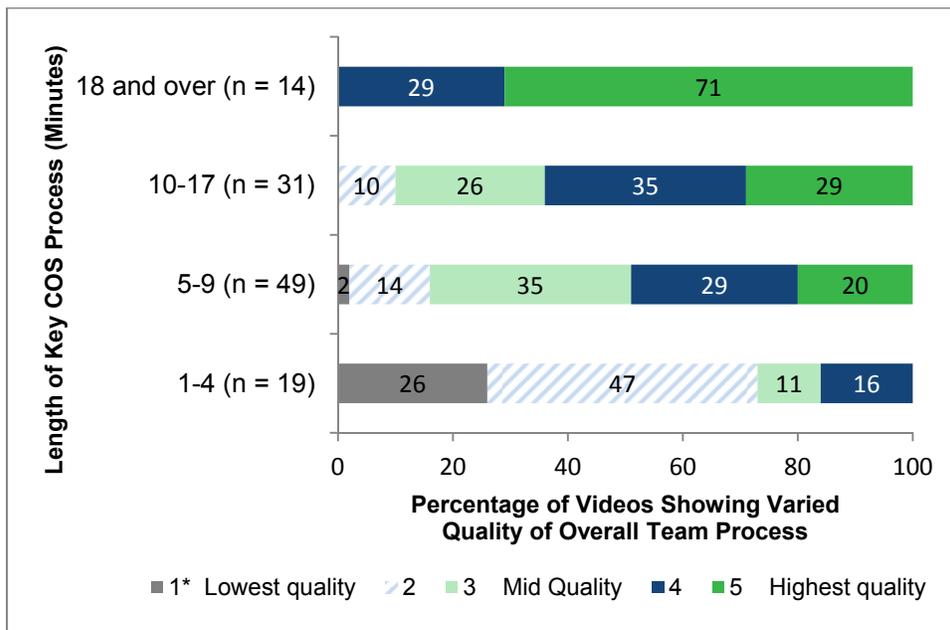
* Two videos were coded as "Can't tell" for the quality of their overall team process. Given that the team process was so minimal or poor that it was not visible enough to code, these two videos were clustered with the lowest quality group.

During team meetings, variation was observed in the amount and type of input that was shared. An expectation in the COS process guidance is that family input should be incorporated into the discussion, either with a family member as an active participant or by a provider sharing information gathered from the family in other ways so that it is represented at the meeting.

- In the videos, a considerable amount of family input was shared in 46% of meetings.
- Only minimal family input was shared in 49% of meetings.
- In only 5% of meetings, no parent input at all was shared.

Having the parent present for the meeting significantly increased the amount of parent input provided ($p < .05$), but perhaps not as much as one might expect.

- Of the meetings where a parent was present, in 54% a considerable amount of parent input was shared while 46% had minimal parental input.
- In no meetings with a parent present did the parent have no input.
- In the 54 meetings without parents present, 37%, 51%, and 11%, respectively, had sharing of considerable, minimal, and no parent input.
- No statistically significant differences were observed in the level of parent input based on EI versus ECSE or entry versus exit meetings.

127

The coder also rated the dialogue between professionals.

- In 11% of the meetings, very minimal or no dialogue occurred between providers, even though more than one provider was present.
- In 50% of meetings, the dialogue was rated as fully appropriate among all the professionals.
- No statistically significant differences were found between entry and exit meetings.
- Statistically significant differences were observed between EI and ECSE teams ($p < .05$).
  - Among both EI and ECSE teams, about half the teams exhibited dialogue between professionals at fully appropriate levels.
  - 20% of ECSE teams compared with 5% of EI teams were observed to have very minimal or no dialogue between team members.
  - The remainder had some midlevel of dialogue.

Team Fidelity to COS Guidance During COS Process

Each state implementing the COS process also should implement training and quality assurance, but state EI and ECSE programs have varied approaches to ensuring and providers gather high-quality data. Many guidance resources about the COS process have been developed to support states in professional development on the COS process. For example, the Early Childhood Technical Assistance Center (ECTA) website (http://www.ectacenter.org/eco/index.asp) has a professional development section with training materials designed to share guidance, explain key concepts, and practice skills. Technical assistance staff members at ECTA also actively work with state staff upon request to support COS implementation and to promote local programs' ability to collect quality child outcomes data.

This section presents findings about the extent to which some of the quality features encouraged in COS resources were evident in the videos. Specifically, data from videos are presented indicating whether the videos show that teams were

- Demonstrating understanding of the child outcomes
- Considering functional skills across settings and situations
- Referencing relevant information during the COS process
- Demonstrating effective age anchoring
- Demonstrating understanding of COS rating criteria.

Data are also presented about

- Avoiding any explicit intent to alter ratings
- Engaging in what appears overall to be a quality COS process.

Demonstrating Understanding of the Child Outcomes

Because teams identify ratings about child functioning for three outcome areas rather than by traditional developmental domains, it is important to consider whether teams focus on the appropriate skills for each outcome during the COS process.

COS Discussion Categorizes Skills into Outcome Areas

As videos were examined, coders looked for major errors in categorizing children's skills and behaviors into outcome areas during the team's COS discussion. For almost all videos, no major errors in categorizing skills into the three outcomes were observed (Exhibit 5.13).

Errors include discussing categories of skills with the wrong outcome or examples of skills with the wrong outcome in a way that could be influential for the rating. For example, the team might discuss feeding and self-care skills while discussing the knowledge and skills outcome area instead of when discussing action to meet needs. Or the team might point to the child's initiative to verbally express what she wants under positive social relationships instead of under action to meet needs. Minor deviations did not count as errors as long as there was no indication that the minor deviation influenced the mix of skills the team focused on for the rating . Examples of a minor deviation are a provider pointing to some of the following examples for positive social relationships: identifying facial expressions from on flash cards or the child knowing her name, gender, and address. Additional information about errors in classification of skills into the child outcome areas as observed in documentation on the COS form is found in Exhibit 5.42 in the documentation section of results from this study.

Exhibit 5.13   Extent That Outcomes Categorization Was Without Major Errors During Video of COS Process (*n* = 113)

	Positive Social Relationships *n* (%)	Knowledge and Skills *n* (%)	Action to Meet Needs *n* (%)
Number with no major errors categorizing skills in team discussion of outcome	106 (94)	106 (94)	108 (96)
Number with major errors	2 (2)	5 (4)	3 (2)
Number of videos with no description of skills	5 (4)	2 (2)	2 (2)

COS Discussion Addresses the Breadth of the Outcome Areas

As teams synthesize information about a child's functioning to decide the appropriate rating, it is important that they consider the breadth of the skills and behaviors that are relevant to the outcome area. For instance, in the case of positive social relationships, teams should consider the child's relationships with familiar and unfamiliar adults, with other children, and following group rules, and they often also discuss the child's social regulation of emotions or feelings and use of language to engage and sustain social interactions with others. Omitting aspects of an

outcome can mean that teams may not consider all of a child's strengths or challenges in the outcome area and thus assign a rating that is inconsistent with the child's true functioning.

Exhibit 5.14 shows the extent to which the teams on videos considered the breadth of each outcome in their COS discussions. For each outcome, there were a few videos on which specific skills were not described in the discussion. These videos were excluded from the percentages shown here because it was impossible to determine the breadth of skills that the teams considered in their appraisals. Instead, those teams made global statements about how a child was functioning without speaking about anything specific the child was doing.

The findings were the following:

- Most COS discussions did show good or moderate breadth of discussion about the three outcomes.
- A statistically significant relationship was found between the length of time teams spent on the COS and the breadth of the COS discussion about each outcome (all $p < .01$) such that longer meetings allowed more time for good breadth in discussions.
- Distinctions in meeting length were more apparent between teams with discussions of moderate versus good breadth (all $p < .01$) or very limited versus good breadth (all $p < .01$) than between teams with discussions of very limited versus moderate breadth (all $p > .28$) .
- It appears that teams discussed earlier outcomes in more breadth than later ones. It is unclear if this is an order effect or if it is easier to discuss some outcomes in more breadth than others.
- No statistically significant relationships were observed in breadth of discussion when comparing EI versus ECSE programs, discussions at program entry versus program exit meetings, or whether or not a parent was present at the team meeting.

Exhibit 5.14   Extent That Team Discussion Involved Full Breadth of the Outcomes in COS Rating Decision

	Positive Social Relationships *n* (%)	Knowledge and Skills *n* (%)	Action to Meet Needs *n* (%)
Number of videos with skills discussed	110	109	108
Good breadth (key outcome areas, broad enough for good decision)	60 (55)	48 (44)	43 (40)
Moderate breadth (missing one or more key areas; mixed)	40 (36)	45 (41)	51 (47)
Limited breadth (focused almost entirely on one aspect of the outcome area)	10 (9)	16 (15)	14 (13)

Exhibit 5.15 shows the consistency with which team discussions reflected good breadth on each of the outcomes.

- Nearly two-thirds (64%) of teams demonstrated good breadth in their discussions on at least one outcome. Relatively few (19%) demonstrated good breadth of discussion across all three outcomes.

- About one-quarter of teams (27%) discussed one or more of the outcomes in very limited breadth.

- Incorporating the breadth of all three outcomes effectively was not related to EI versus ECSE services, entry or exit meetings, or whether or not a parent or family member was present at the team meeting.

Exhibit 5.15   Breadth of Team Discussions Across All Three Outcome Areas
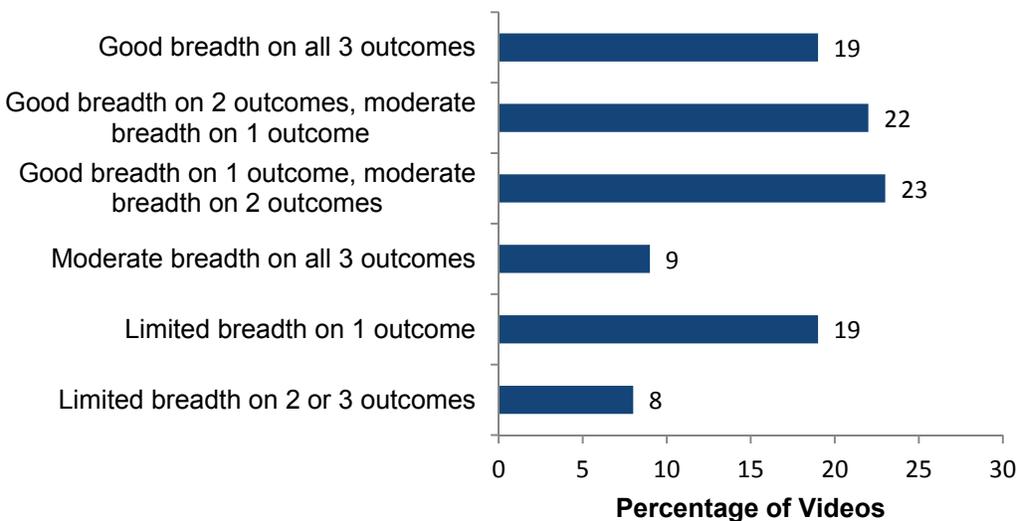
COS Discussion Reflects Depth About the Child's Skills and Functioning

In addition to variation in whether the outcomes discussion focused on the full breadth (or all the aspects) of each outcome, teams varied in the depth of their discussions about skills and behaviors. Some teams discussed the child's functioning in appropriate depth so that all present gained a good sense of how the child functioned with regard to the skills being discussed. Other teams had discussions of mixed depth; they often described how the child used skills or functions with regard to one aspect of the outcome but relied more on global statements or cursory summaries on other aspects discussed.

Exhibit 5.16 shows the depth of discussions about each of the outcomes observed on the videos. A slightly lower percentage of the videos showed discussions of appropriate depth on each outcome than good breadth. Also, the videos showed teams more commonly having cursory or global discussions about the child's functioning than having discussions of limited breadth.

- Appropriate depth was observed across all three outcomes in 25 videos (22%).
- In 47 videos (41%), we observed discussions of cursory depth or global discussion on at least one of the three outcomes.
- A statistically significant relationship was found between the length of time teams spent on the COS and the depth of the COS discussion for each outcome (all , $p < .01$) such that longer meetings allowed more time for appropriate depth in discussions.
  - o For positive social relationships, distinctions in meeting length were more apparent between teams with discussions of cursory versus appropriate depth ($p < .001$) and cursory versus mixed depth ($p < .02$) than between mixed and appropriate depth ($p > .10$).
  - o For the other two outcome areas, distinctions in meeting length were more apparent between teams with discussions that were of mixed versus appropriate depth (both $p < .01$) or of cursory versus appropriate depth (both $p < .01$) than between teams with discussions of cursory versus mixed depth (both $p > .38$).
- No statistically significant relationships were found between the depth of the discussion in videos from EI versus ECSE programs, entry versus exit meetings, or whether or not the parent was included in the COS decision-making discussion.

Exhibit 5.16   Depth of Team Discussions of Skills for Each Outcome in Deciding on COS Rating (n = 113)

	Positive Social Relationships n (%)	Knowledge and Skills n (%)	Action to Meet Needs n (%)
Appropriate depth (good sense of the child's skills about any aspect discussed)	52 (46)	42 (37)	51 (45)
Mixed depth (one aspect discussed with detail about the child's functioning, other aspects discussed briefly with very limited information)	41 (36)	46 (41)	40 (35)
Cursory depth or only global discussion (limited information on how the child functions across the skills discussed)	20 (18)	25 (22)	22 (19)

Depth and breadth are considered independently because the same teams did not always demonstrate both skills. For instance, some teams touched on many different areas (breadth) but shared limited information about those areas (depth). Other teams discussed limited different aspects of the outcome (breadth) but provided much greater detail about how the child functioned with regard to the areas being discussed.

Exhibit 5.17 shows, by outcome area, the percentage of videos in which the team described the child's functioning with depth and addressed the full breadth of the outcome. About one-third of the videos demonstrate both breadth and depth on each outcome.

Exhibit 5.17   Extent Teams Considered Breadth and Depth, by Outcome (*n* = 109)
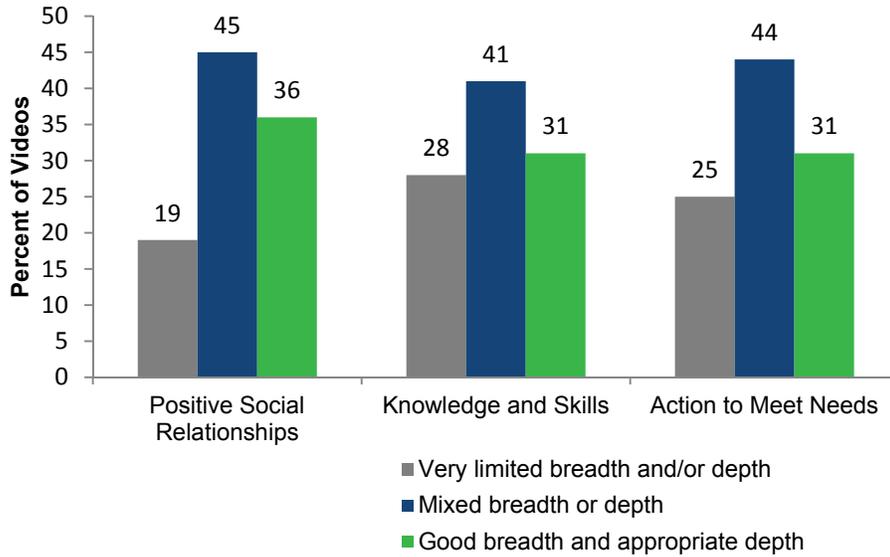
Exhibit 5.18 shows the extent to which appropriate breadth and depth were observed across the three outcomes in the same video. About one-third of videos were coded as having appropriate breadth and depth on at least two of the three outcomes and no outcomes with limited breadth or depth. About one-fifth of the videos had limited breadth and/or depth on two of the three outcome areas.

Exhibit 5.18   Extent of Breadth and Depth in Team Discussions Across All Three Outcome Areas (*n* = 111)
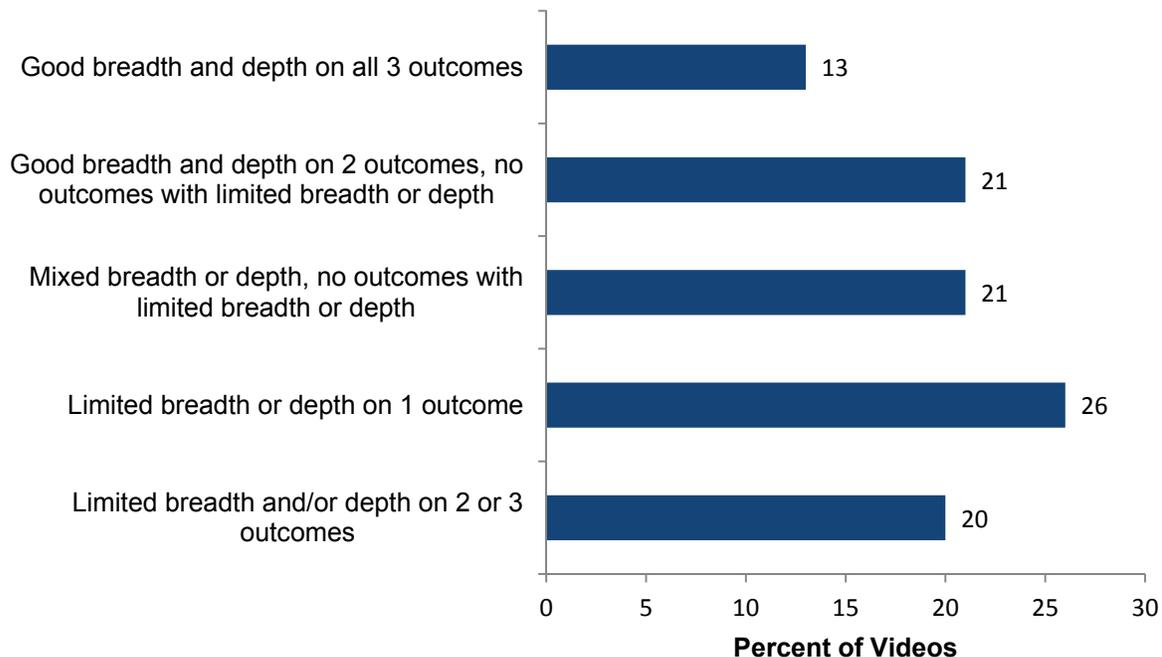
Considering Functional Skills Across Settings and Situations

Effective age anchoring requires providers to identify both the functional skills that the child uses and those that the child does not yet use and then consider where in the sequence of development these skills usually emerge in normative populations. Teams varied in the extent to which the functional skills and behaviors they discussed as part of the COS rating decision were those that represented meaningful behavior in everyday environments on which ratings are expected to be made.

Exhibit 5.19 shows the percentage of videos in which teams mostly referenced functional skills in their descriptions and those in which teams made more limited references to functional skills or did not emphasize functional skills at all.

- Emphasis on functional skills varied across outcomes. A greater percentage of teams described functional skills in discussions about the action to meet needs outcome (61%) than for knowledge and skills (47%), with positive social relationships (58%) in between.
- Most teams described some mix of functional and discrete skills across outcomes, with only 29% of teams referencing mostly functional skills for all three outcome areas and only 2% of teams not referencing functional skills in discussions of any of the outcomes.

Exhibit 5.19   Extent That Teams Referenced Functional Skills in Team Decision-Making Videos (*n* = 113)
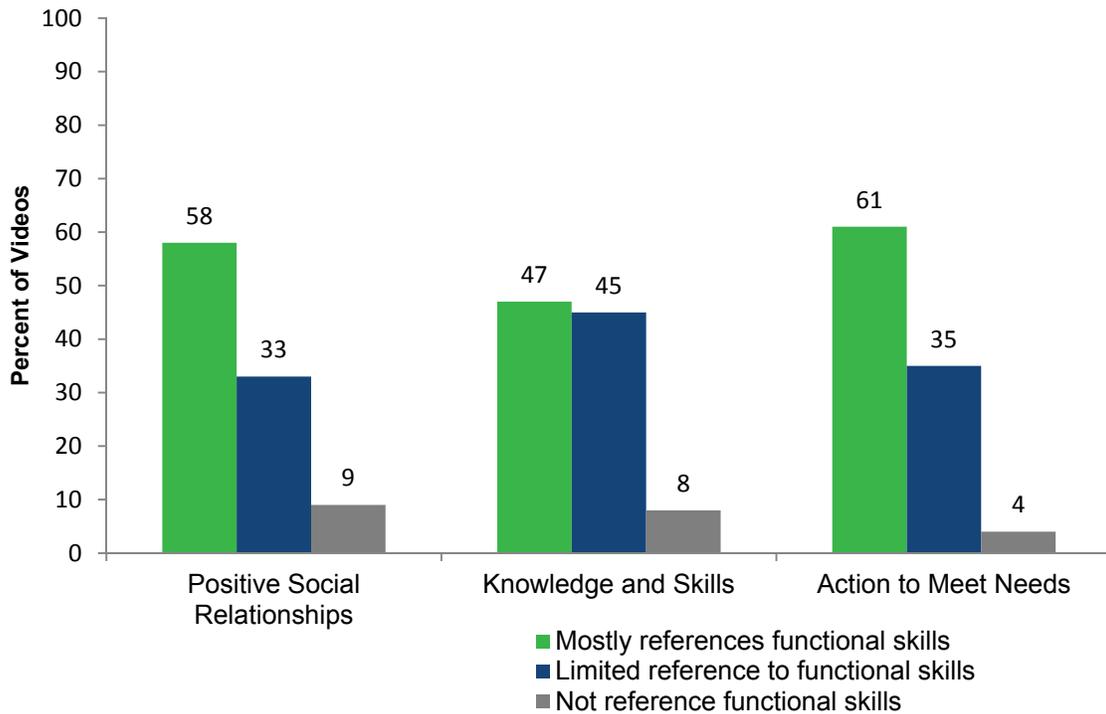
To fully consider the range of the child's skills, it is important for teams to also reference skills that the child does not yet use. Identifying skills the child does not yet use in everyday situations helps gauge the ceiling on how the child uses his or her skills.

- In most videos, teams did reference some skills that the child did not yet use, with 68% of teams doing so in both videos of discussions and on the documentation.
- An additional 21% described skills the child does not yet use in the video but did not record any of them on the documentation.

It also is important that the skills and behaviors discussed reflect the child's functioning across the variety of everyday settings and situations relevant for that child. So, for instance, the team should consider the child's functioning across a variety of situations. Some examples include at home, at child care and/or preschool, with familiar and unfamiliar people, with peers and siblings, during a structured assessment situation, or on playgrounds and in community locations such as restaurants or church nurseries. The specific number of settings to consider varies depending on the everyday experiences of the child, but discussion about multiple settings is warranted in each case.

Exhibit 5.20 shows the number of settings discussed in the videos and whether the number of settings discussed was appropriate for the child.

- For all three outcomes, most teams discussed skills and behaviors in more than one setting.

- Across all three outcomes, teams discussed skills and behaviors in an average of 2.6 different settings.

- When teams described children's positive social relationships, they spoke about significantly more distinct settings and situations than when they talked about children's knowledge and skills (p < .0001) or action to meet needs (p < .0001).

- No statistically significant differences were found in the number of settings discussed by teams when deciding ratings for action to meet needs and knowledge and skills.

The number of settings discussed was related to meeting length but not to EI versus ECSE program type, parent presence at the COS meeting, or entry versus exit meetings. More settings were discussed in longer COS process meetings for positive social relationships (r = .48, p < .0001), knowledge and skills (r = .38, p < .0001), and action to meet needs (r = .42, p < .0001).

Exhibit 5.20   Number of Settings Considered During COS Decision-Making Discussions (*n* = 113)

	Positive Social Relationships *n* (%)	Knowledge and Skills *n* (%)	Action to Meet Needs *n* (%)
Number of settings discussed - Mean (*SD*)	3.1 (1.4)	2.4 (1.1)	2.5 (1.0)
1	15 (13)	20 (18)	13 (12)
2	27 (24)	52 (46)	51 (45)
3	24 (21)	23 (20)	34 (30)
4	35 (31)	15 (13)	10 (9)
5–7	12 (11)	3 (3)	5 (4)
Average number of settings discussed across all three outcomes	*M* = 2.6 *SD* = 1.0		
Less than 2	19 (17)		
2–2.9	46 (41)		
3–3.9	36 (32)		
4–6	12 (11)		
Number of settings discussed seems appropriate for the child	81 (72)	60 (53)	71 (63)
Settings discussed were…			
Appropriate for child on all three outcomes	46 (41)		
Appropriate on two of the outcomes	24 (21)		
Not appropriate on two of the outcomes	26 (23)		
Not appropriate for child on any of the three outcomes	17 (15)		

Referencing Relevant Information During the COS Process

Assessment Information

Throughout the decision-making process, teams discussed their observations and shared information about the child's functioning. Teams based their decisions on information from both formal and informal information sources during the COS process. Informal sources of information relied heavily on observations from service providers, parents, and child care providers. Each of these sources was familiar with the child's functioning in different circumstances or settings. As people shared their informal observations, they often incorporated their opinion of the child's developmental status and helped the group make a rating decision. Sharing informal information often led to rich conversations among team members and helped the group reach a decision about the appropriate rating for the child's level of functioning.

Formal assessment data included scores or information on specific skills from a variety of direct and authentic assessments designed to capture early childhood development across developmental domains. Exhibit 5.21 shows how extensively teams referenced one or more formal assessment tools on their documentation, during the video of team discussion, or both. It also identifies the type of information referenced in the video or on the documentation.

- Most teams (61%) referenced an assessment tool in some aspect of the COS process.
- Only 31% of teams both discussed it on the video and documented it on the form.

Some variation was evident across videos in whether and how (i.e., in discussion on video and/or on documentation of COS form) teams referenced specific assessment tools.

- Reference to a specific assessment tool was more likely to occur in ECSE videos and documentation than in EI videos and documentation ($p < .001$).
- No statistically significant differences were observed between entry and exit meetings.
- A trend was found that teams were somewhat less likely to reference a specific assessment tool ($p = .067$) when the parent was present at the meeting than when the parent was not. Although nearly the same number referenced an assessment tool in some way, in 60% of those meetings where parents were present teams referenced the assessment tool only on their documentation (i.e., did not discuss it on the video), whereas in only 9% of meetings without a parent present teams that referenced an assessment in some way referenced it only on their documentation.

Exhibit 5.21   Information About Assessment Tool Use During the COS Process (n = 113)

	Entire Sample *n* (%)
References/names specific assessment tool	
In video and documentation	35 (31)
In video only	10 (9)
On documentation only	24 (21)
Neither video nor documentation reference specific assessment tools	44 (39)
Scores or age levels from an assessment tool mentioned in video	44 (39)
Scores or age levels from an assessment tool mentioned on documentation	40 (35)
Specific assessment tool content other than scores or age levels mentioned in video	
Many instances	21 (19)
Some instance(s)	32 (28)
No specific content mentioned	60 (53)

The COS process was designed in part to enable teams to use information from diverse sources, including a variety of assessment tools, to inform ratings about the child's functioning.

- The most common assessment tools[25] mentioned either in the documentation or in the video were Creative Curriculum (16), Battelle Developmental Inventory (BDI, 14), Infant-Toddler Developmental Assessment (IDA, 12), and the Preschool Language Scale (PLS, 10).

- Less frequently mentioned assessment tools were DAYC (7); Hawaii Early Learning Profile (HELP, 5); Transdisciplinary Play-Based Assessment (TPBA, 5); Ages and Stages (ASQ) or ASQ-Social Emotional Questionnaire (5); Adaptive Behavior Assessment System (ABAS, 4); Assessment, Evaluation, and Programming System (AEPS, 3); Rosetti (3); Wechsler Preschool and Primary Scale of Intelligence (WPPSI, 3); Sensory Profile (3); Receptive-Expressive Emergent Language (REEL, 3); and Differential Abilities Scale (DAS, 3).

- A total of 25 other assessment tools were mentioned in only one or two instances.

- References to assessment tools during team discussions usually were embedded as part of the natural flow of the description of the child's functioning for each outcome; teams did not have a specific part of the conversation exclusively on reviewing assessment scores or results.

Reviewing Entry Ratings at Exit Meetings

Although reviewing entry ratings at exit is not necessarily a problem for quality ratings, it is recommended that the team complete entry and exit ratings independently in case reviewing earlier ratings for the child evokes bias in the final decision making. Review of team decision-making videos revealed the following:

- In 20% of the videos of exit meetings, the team mentioned the child's COS rating at entry.

- In all these instances, the team discussed the entry ratings only for positive social relationships, not for any of the other outcome areas.

- Discussions about entry ratings were significantly more likely to occur in ECSE videos than EI videos ($p < .05$).

- Discussion about entry ratings at exit was not related to having a parent or family member participate on the team.

COS Decision Tree

The decision tree is a resource developed by the ECO Center that is available to states to support COS decision making. It identifies key questions and criteria that differentiate between the 7 rating points on the COS. Many states encourage staff to use the decision tree during COS discussions.

---

[25] All versions of the assessment tool indicated are counted together. In cases where both a full developmental version and a screening tool are available, the two are collapsed.

- Widespread use of the decision tree was observed in the team decision-making study; overall, 80% of teams referenced the decision tree during the COS meeting, with 60% demonstrating considerable use of it and another 20% acknowledging it or referring to it in some way.
- The decision tree was used more frequently by ECSE teams than EI teams ($p < .01$) and in longer COS meetings ($p < .05$). Exhibit 5.22 presents usage patterns among EI and ECSE programs.
- No statistically significant differences were observed in use of the decision tree at entry versus exit meetings or when parents/family members were present or absent from the COS meeting.

Exhibit 5.22   Use of the Decision Tree Resource for the COS Decision Among Programs ($n = 112$)

Legend:
- Not Used
- Minimal Reference to Decision Tree
- Considerable Use of Decision Tree
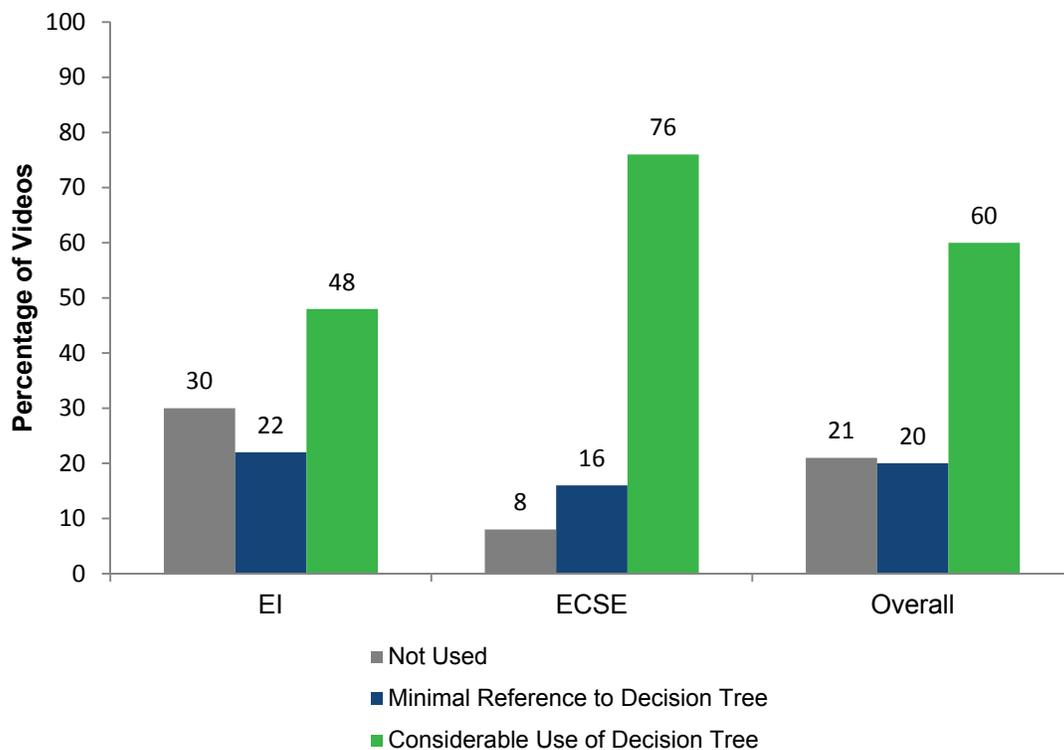
Demonstrating Effective Age Anchoring

In order for team COS ratings to accurately reflect children's true levels of functioning in each of the three outcomes, EI and ECSE providers need to have knowledge about sequences of child development and age-expected skills and behaviors during early childhood, and they need to be able to identify how the child's functioning relates to those age-expected skills and behaviors. Without such knowledge among people involved in the COS decision-making process, teams will have difficulty identifying COS ratings that meet the expected rating criteria.

We examined the videos for explicit age-anchoring references. That is, we looked for instances where teams described a specific skill or behavior of the child and provided some indication of the level of functioning that the skill suggested. Teams tended to examine the child's functional level using some shorthand references and fluid conversational approaches to summary judgments. So, for example, teams did not list 20 skills and provide specific ages when those skills are expected to occur and the age when they began being observed for this particular child. Instead, most teams had a conversation that described the child's functioning and made conclusions about how this functioning related to age-expected levels.

Most teams did find some way to discuss their perceptions about the level of the child's functioning.

- In most videos (81%), teams referenced rating criteria language,[26] using the terms *age expected* or *age appropriate*, *immediate foundational*, or *foundational* in their discussions about functioning in an outcome area. These terms have been used in COS professional development activities; they are linked to specific rating criteria and provide a broad sense about how the child's functioning relates to what is expected among children that age.
- In about half the videos (51%), teams provided a sense of the child's functioning by mentioning specific sequences in the progression that skill development or specific ages when talking about a skill.

Teams varied in how detailed their references to specific skills were during discussions. Consistent with findings about the depth of discussions, in 26 videos (23%) there was not enough detail about the child's functioning on specific skills or the team's appraisal of how they gauged the age level of that skill to determine whether the team age anchored effectively.

Exhibit 5.23 shows the extent to which teams age anchored skills effectively in the videos in which there was enough information about the skill being referenced and the team's appraisal of the age level or corresponding rating criteria language associated with that skill given the child's age.

Exhibit 5.23   Extent of Errors in Age Anchoring in Team Decision-Making Study Videos with Enough Detail to Evaluate Age Anchoring (*n* = 87)

Errors in Age Anchoring (Across All Outcomes)	Subset with Any Age Anchoring *n* (%)
No errors	73 (84)
Minor errors	6 (7)
Major errors	8 (9)

---

[26] This is the same language used on the decision tree resource document.

141

Age-anchoring errors were coded when a team member appraised a skill at a wrong age range or out of sequence in the order it emerges in child development and the observation was not corrected by another member of the team. Another fairly common type of error was for teams to indicate that a skill was age expected when it was indeed a skill that typically emerges much earlier (e.g., walking in a four-year-old); in situations like these, where skills emerge much earlier, the skill should not be categorized as evidence that the child is functioning at age-expected levels. Any error across discussion of the skills on all three outcomes was counted. Errors were coded as major if the misappraisal was notable enough that it would shift a skill across rating categories (e.g., from viewing functioning on that skill from immediate foundational to age expected because the age anchoring was off by a number of months), especially if it was the only skill pointed to with a specific level of functioning so that it would influence the mix of functioning and the likely rating the child would receive. An example of a minor error is misattributing the age when a skill usually emerges so that both the actual skill and the mistaken timing point to the skill being at a foundational level for the child's age, and it thus does not influence the team's appraisal of the child's functioning relative to the rating criteria.

In addition to looking at how teams anchor specific skills discussed, coders reviewed videos for misunderstandings or misapplications of rating criteria. More information about those misunderstandings is described below, but one kind of misapplication is described here. As teams applied the rating criteria to the child's functioning, they sometimes categorized the rating as being on the wrong side of the decision tree or had confusion about or made errors in determining which of neighboring numbers should apply to the child. These kinds of major misunderstandings or misapplications could sometimes be attributed to team errors in thinking about sequences of skill development or in age-anchoring skills. For instance, because a team viewed walking as an age-expected skill for a four-year-old, it referenced the wrong branch of the decision tree and did not consider ratings of 1–3 for the child at all. The error was not in understanding the rating criteria but in identifying the child's skill as age expected.

Exhibit 5.24 shows the percentage of videos in which we observed misunderstandings or misapplications of the rating criteria with a major impact as a result of age anchoring or sequencing problems. Data are shown for each outcome.

Exhibit 5.24   Extent That Age Anchoring Influenced Problematic Application of Rating Criteria, by Outcome (*n* = 113)

	Positive Social Relationships *n* (%)	Knowledge and Skills *n* (%)	Action to Meet Needs *n* (%)
Poor application of rating criteria with major impact due to age anchoring or sequencing problems	11 (10)	2 (2)	0 (0)
Videos without rating criteria problems based on age anchoring only	101 (89)	111 (98)	113 (100)

Teams had more difficulty applying rating criteria because of age-anchoring issues for positive social relationships than for the other outcome areas. This finding is not surprising given that many early childhood programs have more limited availability of effective assessment information about children's social functioning and somewhat less emphasis on skill levels as part of most eligibility determinations. While effective age anchoring is important to implement the COS process effectively, child development knowledge is a skill that providers are expected to have as a necessary foundation for service provision.

Demonstrating Understanding of COS Rating Criteria

Effective ratings require teams to understand and implement the COS rating criteria. Below, we describe data from the team decision-making study about misapplications of the rating criteria. We also summarize some of the issues that contributed to challenges with effective COS ratings.

Frequency and Types of Misunderstandings and Misapplications of COS Rating Criteria

Exhibit 5.25 shows the extent of misunderstanding/misapplication of COS rating criteria. Videos were coded for presence of misunderstandings and their consequences. The exhibit shows misunderstandings where the challenge(s) led to a wrong answer that caused the team to categorize the child's functioning down the wrong pathway of the decision tree or led to confusions and/or errors between neighboring rating criteria numbers.[27]

- Most misapplication of rating criteria occurred with positive social relationships.
- Despite variable amounts of training among team members, in more than 85% of the videos no major misapplications of rating criteria occurred for two of the three outcomes.
- Misapplication of the rating criteria (other than age anchoring) occurred on any of the three outcomes in 40% of the videos.

---

[27] In some cases where misunderstandings are counted, teams did subsequently discuss and correct for problems in that outcome or for interpreting other outcome areas. However, in each of those counted in Exhibit 5.25, during a portion of the discussion the team considered applying a rating that would have been an error based on a misunderstanding or misapplication of rating criteria.

Exhibit 5.25 Extent of Problematic Application of Rating Criteria, by Outcome, on Videos in Which Misunderstandings Led to a Major Influence on the Rating Decision (*n* = 113)

	Positive Social Relationships *n* (%)	Knowledge and Skills *n* (%)	Action to Meet Needs *n* (%)	On Any of the Three Outcomes *n* (%)
Misunderstanding about/misapplication of rating criteria for reasons other than age anchoring (major influence on rating approach)	26 (23%)	14 (12%)	16 (14%)	45 (40%)
Videos without major rating criteria misunderstandings/misapplications for reasons other than age anchoring	87 (77%)	99 (88%)	97 (86%)	69 (60%)

Different issues can underlie problems in applying rating criteria appropriately, and sometimes multiple issues occurred within the same video. Coders categorized problems that were observed in the videos for each outcome (Exhibit 5.26).

- In 1–3% of videos, teams simply had a wrong understanding of the criteria, labeling the meaning of a criterion wrong or interpreting critical decision points differently than the guidance specifies.

- For positive social relationships and knowledge and skills, 3–4% of videos indicated problems because the teams weighted one particular skill or aspect of an outcome too strongly.[28]

- The most common misapplication of rating criteria at exit occurred as teams decided on a specific rating based on the progress a child made from one time point to the next rather than based on the status of the child's functioning relative to what was expected of a child that age. This misunderstanding was observed more often when teams discussed positive social relationships (11%) than knowledge and skills (5%) or action to meet needs (5%).

- Finally, some teams had other issues underlying problems with rating criteria. For instance, teams might have rated a child's functioning against a lower threshold of expectations because of the type of disability the child had, or less clear-cut problems occurred such as different ratings decided by the team verbally than were recorded on the documentation form or teams describing a rationale for a given rating but assigning a number that was not associated with that rationale. Once again, these other types of misunderstandings were more common with the discussions about positive social relationships (9%) than with knowledge and skills (1%) or action to meet needs (5%).

---

[28] Ratings are expected to be based on how the child functions across his or her everyday environments. Sometimes a skill has been observed only in one isolated instance and/or has been observed only in one unique setting. Teams can err in weighting these isolated skills or instances too heavily in the mix of skills to determine ratings across an entire outcome area.

Exhibit 5.26   Types Misunderstandings or Misapplications of Rating Criteria with Major Impact on Ratings by Outcome (*n* = 113)

Type of Misunderstanding or Misapplication of Rating Criteria	Positive Social Relationships *n* (%)	Knowledge and Skills *n* (%)	Action to Meet Needs *n* (%)
Applying criteria the wrong way	1 (1)	2 (2)	3 (3)
Overemphasizing one aspect of rating (e.g., EVER)	3 (3)	5 (4)	1 (1)
Basing rating on child's progress instead of status	12 (11)	6 (5)	6 (5)
Any other poor application of rating criteria or misunderstanding	10 (9)	1 (1)	6 (5)

Team Decision-Making in Applying Rating Criteria

Most teams considered more than one possible rating when they were applying rating criteria; this was more common for knowledge and skills and action to meet needs (76% and 74% respectively) than for positive social relationships (64%).

Often, one or more team members described a rationale for the rating selected as part of the team discussion and checked for consensus. Exhibit 5.27 shows the frequency of this approach and the extent to which verbal rationales were consistent with, or sufficient to justify, the rating that the team finally selected.

- In almost all videos, team members articulated the rationale for the chosen ratings.
- In most cases, the rationales for the ratings were consistent with the chosen ratings.
- About half the discussions about positive social relationships and two-thirds of the discussions about knowledge and skills and action to meet needs included rationales that were sufficient to justify the rating selected. Many of the rationales could not be identified as wholly sufficient to justify the rating because the teams did not explicitly state the presence or absence of skills a child had in a way that gave the coder confidence that the team had considered all aspects of the rating criteria before deciding on the rating.
- None of the videos showed teams that had unresolved disagreements at the end of the COS process meeting; all the teams managed to reach consensus in some way on the rating. This does not mean that all team members were asked to verbally indicate their agreement with the rating, but rather that there was no explicit unresolved verbal disagreement about the rating at the end of the video segment on each outcome.

Exhibit 5.27 Rationale Described During COS Discussions

	Positive Social Relationships *n* (%)	Knowledge and Skills *n* (%)	Action to Meet Needs *n* (%)
Team member articulated a rationale for rating (*n* = 113)	104 (92)	99 (88)	99 (88)
Relationship of rationale to rating (*n* = number with rationale on rating)			
Consistent with rating selected	76 (73)	78 (79)	81 (82)
Sufficient to justify rating selected	48 (46)	64 (65)	64 (65)

Evidence About Explicit Intent to Alter Ratings

In very few videos (5 of the 113 videos, 4%) were there instances where one or more team members indicated that the rating selection would make the program look good during the rating discussion. All these were teams that discussed it in the context of inflating exit scores rather than selecting lower entry scores.

Overall Quality of the COS Process

After rating many specific indicators of quality COS practices, coders also rated the overall quality of the COS process on a scale from 1 to 5 for each outcome on the video. Exhibit 5.28 shows the results.

- For each of the three outcomes, over half the videos were rated high quality (4 or 5) for (55%, 60%, 64%, respectively), with another about one-third rated medium quality (3) (33%, 31%, 24%, respectively).

- Notably, although on average teams spent longer discussing positive social relationships and talked about a somewhat greater number of settings in that outcomes discussion, the quality of the COS process was not notably higher for positive social relationships than for other two outcome areas, perhaps because teams also exhibited weaker age anchoring and more problems basing ratings on progress rather than status information with positive social relationships than the other outcomes.

- Mean coder ratings of the quality of the COS process were 3.5 (*SD* = 0.9), 3.6 (*SD* = 0.8), and 3.6 (*SD* = 0.9) for positive social relationships, knowledge and skills, and action to meet needs, respectively.

- No statistically significant differences were found in mean ratings when comparing videos from EI and ECSE programs, entry and exit meetings, whether or not the parent was included on the team, or whether the child had a higher or lower score on the ABILITIES Index.

- Specific COS ratings assigned by the team were unrelated to coder ratings of overall COS quality.

- Being in a program where many providers had training and experience with the COS was not related to overall quality of the COS process in the video. There were no statistically significant differences in the overall quality of the COS process in videos from programs where the program had higher percentages of providers who had 31 or more COS process experiences and where they had received 9 or more hours of training on the COS.
- There was statistically significant higher COS quality when all the providers on the child's team indicated that they understood very well the skills and behaviors associated with the three outcomes, the definition for the 7 rating points, and the degree to which different skills and behaviors are age appropriate (high self-rated knowledge $M = 3.7$, SD 0.8; lower self-rated knowledge $M = 3.3$, $SD = 1.0$, $F = -2.11$, $p = .037$).
- Several other factors also were related to higher ratings of overall COS quality:
  o Longer COS discussions were associated with higher quality COS process ratings with correlations of .27, .41, and .27, respectively, for positive social relationships, knowledge and skills, and action to meet needs (all $p < .01$).
  o Higher ratings of team process quality also were associated with higher ratings of a quality COS process on each outcome ($r = .49$, $p < .001$ for positive social relationships, $r = .62$, $p < .001$ for knowledge and skills, $r = .59$, $p < .001$ for action to meet needs).

Exhibit 5.28   Coder Rating of Overall Quality of the COS Process, by Outcome ($n = 113$)
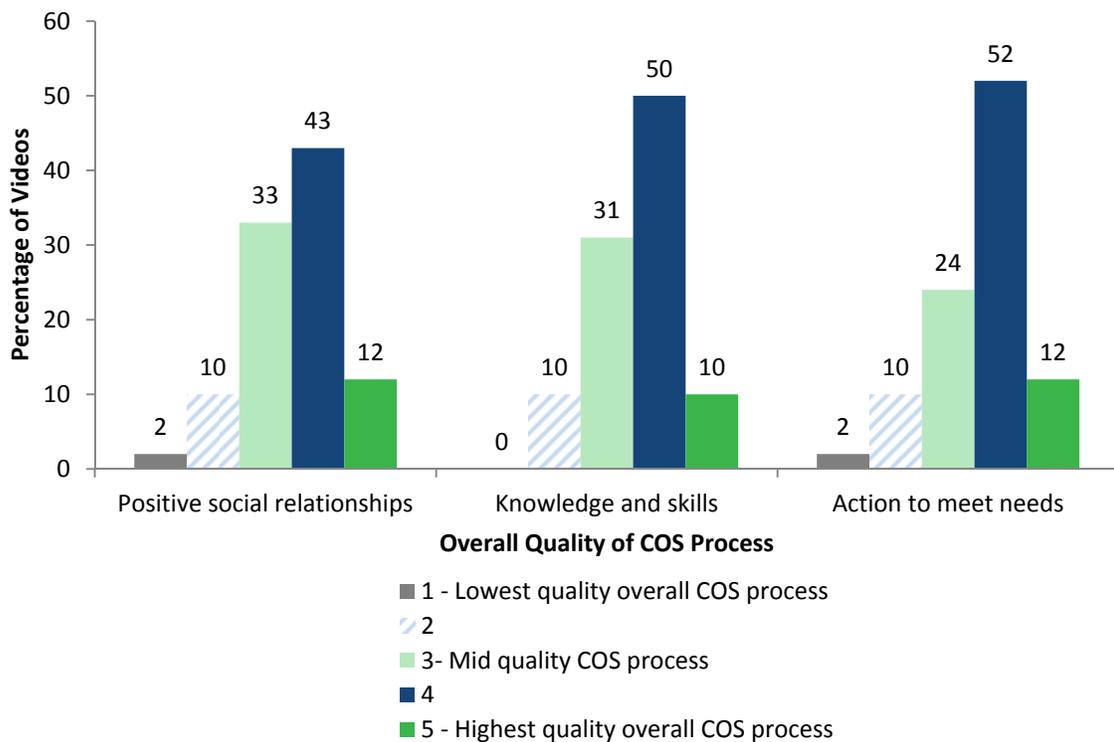
Consistency of Ratings by Teams versus Coders

Methodological Approach

One important question to consider is the extent to which videos showed team COS rating decisions that were consistent with the ratings that would be assigned by those highly trained in understanding the COS rating criteria who had access to the same amount of information about the child. Testing this question is challenging methodologically because accurately identifying a rating requires knowing the child's functioning across settings and situations, usually with input and discussion across team members who have experience with the child in a variety of settings. It is not clear that either the coder or the team represents a gold standard in rating accuracy. A highly-trained, "objective" coder lacks the extensive experience with the child, access to all the information in the child's record about the child, and the shared understanding with other team members that would help in interpreting team members' often abbreviated explanations about the child's functioning or explanation of assessment findings. Additionally, all those on the team who know the child well may or may not have strong understanding of the COS rating criteria and COS process recommendations. Given the inherent challenges of examining matches between an outside rater and an actual team that has worked with the child, the researchers established that ratings within one point of each other would count as a match.

In the team decision-making study, we examined the match in ratings between a highly trained coder watching the video and the team in the video. In some of the videos, teams described the child's functioning well enough for the coder to identify a single rating point as the COS rating. In other videos, however, there was not enough information about a child's functioning for the coder to confidently select a single rating point. Videos might be brief and not contain information of sufficient depth about skills discussed to age anchor them confidently, or teams might talk in shorthand, referencing information shared with each other in earlier meetings or based on shared observations where the coder was not present. Often, coders had enough information to rule out certain ratings and identify a range of two ratings where the rating would fall but not necessarily be confident about a single rating point. Sometimes ranges could be narrowed down only to three or four points on the rating scale. For instance, if the only thing the coder was certain about was that the team agreed the child had not exhibited any age-expected functioning, then the coder knew the rating was between 1 and 3 but not the specific mix of skills or age level of skills that could differentiate between ratings of 1, 2, or 3.

Exhibit 5.29 shows how coder and team ratings were analyzed and classified as matches, not matches, or as having insufficient information to determine whether the ratings matched or not.

- Instances where the team and coder were within one point of each other were counted as matches.
- Instances where the team and coder were two or more points away from each other were counted as *not* being matches.

- Instances where the team and coder might be one or two points away were coded as *can't tell*.
- If coders could not identify either a single point or a two-point range as the appropriate rating, the video was classified as can't tell based on insufficient information.

Exhibit 5.29   Classification Approach for Match Versus Not a Match in Ratings in Team Decision-Making Study

Coder Rating Type	Difference of Team Minus Coder Rating	Example	Classification
Single point	0 or 1	• Coder rating 5, Team rating 5 • Coder rating 4, Team rating 5 • Coder rating 6, Team rating 5	Match
	2-3	• Coder rating 3, Team rating 5 • Coder rating 2 or lower, Team rating 5	Not a match
Range of two numbers	0	• Coder rating 4–5, Team rating 4 • Coder rating 4–5, Team rating 5	Match
	1	• Coder rating 4–5, Team rating 3 • Coder rating 4–5, Team rating 6	Can't tell
	2-3	• Coder rating 4–5, Team rating 2 • Coder rating 4–5, Team rating 7	Not a match
Range of three or more numbers	Any distance	• Coder rating of 1–3, Team rating of 1 • Coder rating of 1–3, Team rating of 3 • Coder rating of 1–3, Team rating of 5	Can't tell

Note: Coder Rating Type: While, as shown, coders varied in how precise their ratings were, all teams in videos ultimately decided on a single point rating for each outcome.

Differences are based on absolute values, since whether the team or the coder rating is higher is not important; simply the distance of these ratings from each other is to be noted.

Match of Team and Coder Ratings

Coders' and teams' ratings matched for a majority of the videos. Exhibits 5-30 and 5-31 show the extent of matching between coders watching the videos and the team rating that was identified.

- Across the 339 outcomes (3 outcomes for each of 113 videos), a definitive *match* or *not match* classification could be made for 255 outcomes (75%).
- Of these videos where a definitive match or not match classification could be made, a match between coder and team ratings was found in 93% of the videos.
- Collapsing across all outcomes, 93% of single-point ratings matched and 92% of two-point range ratings matched.

- When a single-point rating was identified, match percentages were 93%, 88%, and 97% for each of the three outcome areas (positive social relationships, knowledge and skills, and action to meet needs, respectively) and 85%, 97%, and 93%, respectively, when a two-point range was identified.[29]
- In about one-fourth of videos, there was insufficient information about one or more outcomes for the coder to determine a rating to match to the team's rating (i.e., 29%, 24%, 21% of videos respectively, had insufficient information for definitive match information on positive social relationships, knowledge and skills, and action to meet needs).

Exhibit 5.30   Match between Coder and Team Ratings (*n* = 113)

Classification	Positive Social Relationships			Knowledge and Skills			Action to Meet Needs		
		Percentage of			Percentage of			Percentage of	
	n	All Videos	Videos with Infor-mation	*n*	All Videos	Videos with Infor-mation	*n*	All Videos	Videos with Infor-mation
Match	71	63	89	81	72	94	84	74	94
Not match	9	8	11	5	4	6	5	4	6
Can't tell/ excluded based on insufficient information	33	29		27	24		24	21	

---

[29]   These percentages exclude the videos where there was insufficient information to make a definitive match-not match determination.

Exhibit 5.31   Detail on Coder-Team Match on Ratings in Videos

Coder Rating	Classification	Positive Social Relationships			Knowledge and Skills			Action to Meet Needs		
		n	Percent		*n*	Percent		*n*	Percent	
Single point										
Same rating	Match	30	75	93	20	77	88	27	84	97
Within 1 point	Match	7	18		3	12		4	13	
2 or more points away	Not match	3	8		3	12		1	3	
Two-point range										
Within range	Match	34	85		58	97		53	93	
2 or more points away	Not match	6	15		2	3		4	7	
Overall match rate	**Match**	71/80	89		81/86	94		84/89	94	
Can't tell/ excluded based on insufficient information		33	29		27	24		24	21	

Influence of Parent Involvement on COS Process and Ratings

As the COS process began being implemented, there was much discussion nationally and within states on setting policies about whether or not to include parents or family members in the COS discussion and how that might influence ratings. In the team decision-making study, about half (52%) of the videos were of teams with a parent or family member present, and percentages with family members present were similar for EI and ECSE meetings and entry and exit meetings. Data from this study help describe the extent of family member participation in the videos, background information shared with parents when they were present in COS meetings, and the influence of parent or family member participation in the meeting on COS ratings.

COS Process in Meetings Including Parents or Family Members

On meeting videos where background information was shared with parents, the data showed the following (Exhibit 5.32):

- In two-thirds of the videos, some explanation of why outcomes data were being collected was given (69%), and the meaning of each outcome was described (65%).
- In half the videos (49%), no description of skills expected for the child's age was provided.
- Parents were actively engaged in most videos; in 73% of them, parents spontaneously spoke up and shared information.

Exhibit 5.32   Background Information Shared with Parents Present at Meeting About COS Decision (*n* = 59)

Characteristics of Discussion with Parent	*n* (%)
Explanation about why outcomes data are collected	
Explains data are collected for program improvement/ accountability and to observe the child's progress	6 (15)
Explains data are collected for program improvement/ accountability only	5 (12)
Explains data are collected to observe the child's progress only	17 (42)
Explains why data are collected, giving other reason	11 (19)
Does not explain but references earlier conversation about why data are collected	2 (5)
Does not explain at all	18 (31)
Description of meaning of each outcome*	
Describes meaning of each outcome	29 (49)
Describes meaning of some, but not all, outcomes	9 (15)
Does not describe, but references earlier description	1 (2)
Does not describe at all	21 (36)
Description of skills expected for the child's age	
Describes expected skills for each outcome	10 (17)
Describes expected skills for some, but not all, outcomes	12 (20)
Gives broad description of skills for child's age; cannot link to specific outcome	8 (14)
Does not describe at all	29 (49)
Description of idea of sequences in development	8 (14)
Provider encourages parent involvement	
Encourages parent to share or asks open-ended questions	49 (83)
Only asks parent if agrees	10 (17)
Parent engagement in video	
Spontaneously speaks up and shares information	43 (73)
Nods head/agrees only	11 (19)
Shows no spontaneous involvement or sharing	5 (8)

* More than 1 category could be coded among these options

Influence of Parent/Family Involvement on COS Ratings

No statistically significant differences were found between team COS ratings on each of the outcomes for teams when a parent or family member was at the meeting and when no parents or family members were at the meeting (Exhibit 5.33).

Exhibit 5.33   Average Team COS Ratings by Outcome for Videos with and Without Parents at the Meeting (*n* = 113)

	Positive Social Relationships *M* (*SD*)	Knowledge and Skills *M* (*SD*)	Action to Meet Needs *M* (*SD*)	ABILITIES Index Score Reversed *M* (*SD*)
Parent at COS team meeting	5.2 (1.5)	4.8 (1.5)	5.0 (1.5)	102.8 (10.4)
Parent not at COS team meeting	4.8 (1.6)	4.5 (1.6)	4.7 (1.5)	100.7 (12.6)
Statistically significant difference?	No	No	No	No

Influence of Other Factors on the COS Process and Ratings

An important issue for the validity of a tool is to understand how other factors influence the COS ratings. We examined COS ratings in relation to

- severity of the child's impairment,
- child's disability type, and
- other child and team characteristics.

Relationships Between COS Ratings and the Child's Level of Functioning

COS ratings were examined relative to the severity of the child's disability using the ABILITIES Index (Exhibit 5.34). Correlations between COS ratings and the ABILITIES Index scores are shown for both the samples of entry (*n* = 73) and exit (*n* = 40) meetings in the team decision-making study.

- Correlations indicate moderate to strong relationships between the child's overall functioning as measured by the ABILITIES Index and the child's COS rating.

Exhibit 5.34   Correlations Between COS Ratings and ABILITIES Index Total Scores and Subareas at Same Time Point in Team Decision-Making Study

Team Decision-Making Study	Correlations with COS Ratings at Same Time Point		
	Positive Social Relationships	Knowledge and skills	Action to Meet Needs
Entry meetings: ABILITIES Index Total reversed (*n* = 73)	.59****	.65****	.71****
ABILITIES Social/Communication	.72****	.70****	.62****
ABILITIES Cognitive/Communication	.62****	.69****	.58****
ABILTIES Structural Integrity	.31**	.39***	.53****
Exit meetings: ABILITIES Index Total reversed (*n* = 40)	.61****	.66****	.60****
ABILITIES Social/Communication	.72****	.65****	.57***
ABILITIES Cognitive/Communication	.68****	.72****	.60****
ABILTIES Structural Integrity	.29	.41**	.37*

* $p < .05$, ** $p < .01$, *** $p < .001$, **** $p < .0001$.

Note: ABILITIES Index scores have been reversed such that lower scores represent lower functioning and higher scores represent higher functioning. Darker numbers are expected to show relationships with COS ratings. Lighter numbers indicate aspects of the child's functioning that are less directly related to the outcome area.

Mean COS ratings also were compared for the approximately one-third of children who scored below 100 on the ABILITIES Index scale and for others with higher scores (Exhibit 5.35). In this case, lower scores indicated that the child's disability or delay was more severely affecting the child's functioning across a range of areas.

- Children with higher levels of functioning had significantly higher COS ratings on average for each of the three outcomes both in entry meetings and in exit meetings.

Exhibit 5.35   **Mean COS Ratings for Children with ABILITIES Index Scores of Less Than 100 and 100 or Higher in Entry and Exit Team Decision-Making Study Samples**

	Positive Social Relationships		Knowledge and Skills		Action to Meet Needs	
	ABILITIES Index* < 100 COS Mean (SD)	ABILITIES Index 100 or Higher COS Mean (SD)	ABILITIES Index* < 100 COS Mean (SD)	ABILITIES Index 100 or Higher COS Mean (SD)	ABILITIES Index* < 100 COS Mean (SD)	ABILITIES Index 100 or Higher COS Mean (SD)
Entry meetings (total *n* = 73; ABILITIES < 100 *n* = 23, 100 or higher *n* = 50)	3.26 (1.54)	5.46*** (1.22)	3.13 (1.74)	4.94*** (1.19)	3.26 (1.25)	5.04*** (1.11)
Exit meetings (total *n* = 40; ABILITIES < 100 *n* = 11, 100 or higher *n* = 29)	4.55 (1.13)	5.83* 1.26)	4.27 (0.90)	5.55* (1.18)	4.18 (1.17)	6.03*** (1.12)

* *p* < .01, ** *p* < .001, *** *p* < .0001.

Note: ABILITIES Index scores have been reversed such that lower scores represent lower functioning and higher scores represent higher functioning.

Relationships Between COS Ratings and Disability Categories

Children With Speech or Language Impairments

Although the samples were small in the team decision-making study, we examined the data to see whether trends expected for children with various types of disabilities followed predicted patterns. For instance, we expected children in ECSE with a speech or language impairment as their primary disability to have higher mean COS ratings on average than children with other primary disability categories for each of the outcome areas.

As seen in Exhibit 5.36,

- For each outcome, children with speech or language impairments had higher mean COS ratings at entry than children with other primary disabilities.

- Given the small numbers in the sample, these differences reached statistical significance only for positive social relationships (*p* < .05), with a trend observed for differences in ratings for action to meet needs.

Similar findings were evident in a regression predicting entry ratings that also took the child's age at the meeting time into account (Exhibit 5.37).

Exhibit 5.36   Mean COS Ratings on Each Outcome for ECSE Children with Speech or Language Impairment and Children with Other Primary Disabilities at Entry

Primary Disability Group	Positive Social Relationships	Knowledge and Skills	Takes Actions to Meet Needs
	COS Rating Mean (SD)	COS Rating Mean (SD)	COS Rating Mean (SD)
Speech or language impairment (*n* = 12)	5.5* (.8)	5.0 (1.5)	5.3^t(1.4)
Other disabilities (*n* = 23)	4.5 (1.3)	4.5 (1.4)	4.4 (1.2)

^{t} = trend *p* ≤ .09, * *p* < .05, ** *p* < .01, *** *p* < .001.

Exhibit 5.37   Regression of Speech or Language Impairment Predicting COS Ratings by Outcome in Entry ECSE Sample (Speech or Language *n* = 12, Other Types of Disabilities *n* = 23)

Variable	Positive Social Relationships			Knowledge and Skills			Action to Meet Needs		
	B	*SE B*	*B*	*B*	*SE B*	*β*	*B*	*SE B*	*β*
Intercept	4.34***	1.13	.00	3.52*	1.37	.00	4.24*	1.21	.00
Child age at COS meeting	.00	.03	.03	.02	.03	.12	.00	.03	.03
Speech or language impairment	.98*	.43	.38	.55	.52	.18	.82^t	.46	.30
R^2		.14			.05			.09	
F		2.6^t			0.76			1.57	

^{t} = trend *p* ≤ .09, * *p* < .05, ** *p* < .01, *** *p* < .001.

Children Who Providers Identified with Other Primary Disabilities: Autism or Spectrum Disorders

It was expected that children in ECSE with autism or autism spectrum disorders would have lower COS ratings on positive social relationships than children with other primary disability categories for eligibility. Exhibit 5.38 shows the mean differences in entry COS ratings between these two groups.

- There was a trend for children with autism to have lower COS ratings for positive social relationships (*p* < .09), and the autism group had significantly lower COS ratings on the other two outcomes (*p* < .05 and *p* < .01, respectively).

Exhibit 5.38   Mean COS Ratings of ECSE Children with Autism/Spectrum Disorders and Children with Other Primary Disabilities at Entry

Indications of Autism	Positive Social Relationships	Knowledge and Skills	Takes Actions to Meet Needs
	COS Rating Mean (SD)	COS Rating Mean (SD)	COS Rating Mean (SD)
Autism/spectrum (*n* = 5)	3.2[t] (1.8)	3.2* (1.8)	3.2** (1.3)
Other primary eligibility category (*n* = 30)	5.1 (0.9)	4.9 (1.3)	5.0 (1.2)

[t] = trend *p* ≤ .09, * *p* < .05, ** *p* < .01, *** *p* < .001.

Exhibit 5.39 shows a regression predicting COS ratings at entry after controlling for the child's age at the entry.

- Despite the small sample, children in the autism group had lower COS ratings in all three outcomes areas. Differences reached statistical significance for positive social relationships (*p* < .001) and knowledge and skills (*p* < .05).

Exhibit 5.39   Regression of Reported Autism Predicting COS Ratings, by Outcome, in Entry ECSE Sample (Autism *n* = 5, Other Types of Disabilities *n* = 30)

Variable	Positive Social Relationships			Knowledge and Skills			Action to Meet Needs		
	B	*SE B*	*β*	*B*	*SE B*	*β*	*B*	*SE B*	*β*
Intercept	6.11***	1.04	0.00***	4.83***	1.31	0.00***	5.82***	1.14	0.00***
Child age at COS meeting	-.02	.02	-.14	.00	.03	.01	-.02	.03	-.12
Autism	-2.06***	.54	-.58***	-1.69*	.68	-.41*	-1.87	.59	-.50
R^2		.31			.17			.24	
F		7.35**			3.33*			5.04*	

* *p* < .05, ** *p* < .01, *** *p* < .001.

Relationships Between Characteristics of Child and COS Team and COS Entry Ratings

To understand how key characteristics of the child and COS team are related to COS entry ratings, we conducted a series of regressions. In each case, we examined the extent to which each variable predicted entry COS ratings after controlling for the ABILITIES Index scores on the subareas related to the outcome. (For more information about the subareas see measures section of Methods.)

A summary of these results is provided in Exhibit 5.40. Detailed results from these regressions[30] are shown in Appendix M.

These regressions showed the following significant relationships with the ABIILITES Index subarea as a covariate:

- Children in EI had lower COS ratings than children in ECSE for knowledge and skills ($p < .01$).
- Children who entered programs at older ages had higher COS ratings for positive social relationships ($p < .05$) and knowledge and skills ($p < .01$).
- Team composition such as the number of individuals on the team or whether parents or service coordinators were included in the meeting did not significantly predict teams' COS ratings, nor did children's gender or race/ethnicity.

---

[30] Appendix M shows details about regressions both using the subareas of the ABILITIES Index as a covariate and the total score of the ABILITIES Index.

Exhibit 5.40   Summary Table of Findings from a Series of Regressions Predicting COS Ratings at Entry in Team Decision-Making Study Videos of Entry Meetings (*n* = 73)

Predictor Variable	Positive Social Relationships			Knowledge and Skills			Action to Meet Needs		
	B	SE B	β	B	SE B	β	B	SE B	β
Early Intervention program	-0.44	0.27	-0.13	-0.79**	0.26	-0.25**	-0.26	0.29	-0.09
Male	-0.23	0.30	-0.06	0.12	0.31	0.04	0.04	0.31	0.01
Race/ethnicity									
Hispanic	-0.65	0.39	-0.15	-0.57	0.39	-0.13	-0.77	0.40	-0.20
African American/ black	-0.23	0.36	-0.06	-0.28	0.36	-0.07	-0.33	0.37	-0.09
Other race/ethnicity	-0.42	0.46	-0.08	-0.54	0.47	-0.11	-0.63	0.47	-0.14
Child age at entry (months)	0.02*	0.01	0.19*	0.03**	0.01	0.25**	0.00	0.01	0.00
Parent or family member in COS rating decision meeting	0.28	0.27	0.08	0.25	0.28	0.08	0.35	0.28	0.13
Number on COS team	0.04	0.06	0.05	-0.07	0.07	-0.10	-0.06	0.07	-0.09

* $p < .05$, ** $p < .01$, *** $p < .001$.

Note: Values shown are for a series of regressions in which the item in the left column was a predictor of COS entry ratings in a model using the outcome-relevant subareas of ABILITIES Index as a covariate in every model (e.g., social communication for positive social relationships; cognitive/communication for knowledge and skills; structural integrity for action to meet needs).

Documentation of the COS Process on the COS Form

Because states can and do use the completed COS forms for monitoring and quality assurance purposes, we reviewed the documentation on the COS forms in this study. We examined completeness of forms as well as some minimal indicators of fidelity to COS guidance evident from the documentation.

Completeness of Forms

After deciding on a COS rating, teams complete a form documenting their rating decision. The completeness of documentation is important for identifying potential quality assurance issues for COS data in states. We found the following with regard to documentation:

- On all 113 forms, 100% of teams recorded ratings for all three outcomes on the form.
- Of the 40 exit meeting videos, a small percentage of teams left the question about whether the child made any progress since the last rating blank (positive social relationships $n = 2$ [5%], knowledge and skills $n = 1$ [3%], action to meet needs $n = 1$, [3%]).
- Completion of the progress question did not necessarily mean that teams discussed the child's progress verbally in the video. Reviewing the videos for discussions about the child's progress showed
  - o 46% of teams in the exit meeting videos discussed the child's progress for each of the three outcomes.
  - o 23% discussed the child's progress on some, but not all, of the outcomes.
  - o 31% did not discuss the child's progress.
  - o None of the conversations in the coded videos suggested that the team had confusion about what is meant by "any progress" in the second question on the COS form[31]. No evidence of mistaken understanding of progress was observed in the 40 exit videos coded for the team decision-making study.

Another area reviewed on forms was whether or not teams documented evidence along with the COS ratings and, if so, what was the quality of that written evidence. We found the following:

- For about two-thirds of the forms, teams in the video had evidence documented, and in about one-fifth of teams had no documentation of evidence. The remaining approximately 10% of submitted COS forms had some, but quite minimal, evidence documented for the rating (Exhibit 5.41).
- EI programs were more likely to lack evidence than ECSE programs ($p < .0001$ on all three outcomes).

---

[31] Prior to this study, we had heard that some providers erroneously believe a child must make progress relative to same-aged peers in order to count as a "yes" on the progress question. In actuality, guidance indicates that that teams should mark yes if a child has begun to use even one new skill or behavior since the entry rating regardless of the child's functioning relative to same-aged peers.

- Within EI, exit meetings were more likely to include evidence than entry meetings ($p <$ .05 on all three outcomes). No statistically significant difference was observed between entry and exit meetings in ECSE.
- About a third of forms without evidence noted that the evidence was somewhere else (e.g., they had a notation to see an evaluation report or see present levels of development PLOD section of IFSP).

Exhibit 5.41   Extent Evidence Is Documented on Child Outcomes Summary (COS) Form (n = 113)

	Positive Social Relationships *n* (%)	Knowledge and Skills *n* (%)	Action to Meet Needs *n* (%)
Evidence documented (more than minimal)	83 (73)	76 (67)	76 (67)
Minimal evidence documented	9 (8)	15 (13)	15 (13)
No evidence on COS form*	21 (19)	22 (19)	22 (19)

Note: In some cases, teams did not record evidence on the COS form itself, but included a note that referred individuals to see information about the child's functioning contained in the present levels of development (PLOD) section of the IFSP.

Quality of Evidence on COS Form

The research team examined the COS form documentation to see how well the evidence on it supported the ratings. We examined the extent to which documentation provided with each video met two different standards (Exhibit 5.42). At the most basic level, coders identified whether or not the evidence was *consistent* with the rating given. Evidence was deemed consistent if there was nothing documented that ran counter to the criteria for the rating. For example, if a rating of 3 was assigned, the child would not be expected to have any age-expected functioning. So evidence would be consistent if the skills described included skills at the immediate foundational level and no skills at an age-expected level. If skills were documented that were age expected, the evidence would be classified as inconsistent.

The second standard examined was the extent to which the documentation was *sufficient to justify* the COS rating. In this case, evidence recorded generally justified the rating; it showed that some necessary age-appropriate, immediate foundational, or foundational skills (as expected based on the rating) were present; gave some sense of the amount of these skills; and indicated enough information to provide confidence in the rating criteria decisions shown on the decision tree. Evidence that met the standard supported each of the relevant decisions associated with the rating criteria. Thus, for a COS rating of 3, evidence had to indicate that none of the child's functioning was at an age-expected level and had to show use of immediate foundational skills and a mix of those skills that spanned across most or all settings and situations rather than only occasional demonstration of use of them.

To justify a rating, teams had to provide enough information to address breadth of the outcomes. If information was omitted about the child's level of functioning on a core aspect of the outcome, then documentation was usually viewed as consistent with the rating, but not fully sufficient to justify it. Missing information could be enough raise questions about a discrepancy between the evidence and the rating and therefore the rating was not fully justified. As seen in Exhibit 5.42,

- About two-thirds of the videotaped meetings had documentation provided that was consistent with the ratings made for each of the three outcomes (71%, 66%, and 70% for the three outcomes respectively), but only about one-third (35%) had such evidence documented across all three outcomes.

- Documentation was sufficient to justify ratings for one-third to half of the videos (i.e., 55%, 36%, and 45% for the three outcomes, respectively). Only 12% of the forms had documentation that was sufficient to justify all three of the ratings.

- Evidence supporting the rating for the knowledge and skills outcome was the least likely to meet either standard.

One possible explanation for lower percentages in the knowledge and skills area is that teams often documented assessment tool domain scores for the knowledge and skills outcome area. Data from these scores sometimes did not agree with one or more of the rating criteria decisions, and could not be interpreted without additional information and explanation. The findings are consistent with the variability in implementation practices observed throughout the project. At the time of data collection, few states had implemented monitoring practices and data quality audits of COSFs to increase consistency in documentation.

Understanding of the Three Child Outcomes on COS Form Documentation

Coders also examined data on errors in documentation on the COS form in categorizing skills into the three child outcomes. These were recorded as present or absent on the entire form rather than for each outcome. Exhibit 5.42 shows the percentage of forms with major errors in categorizing skills into outcomes.

- While nearly all the documentation was without major errors (97% of those with evidence including skills), 22% of all forms or 27% of those with evidence did have at least one instance of information recorded about a skill with the wrong outcome area.

- Documented misclassifications were usually the same as those observed in the video, such as recording information about skills assessed with a domain-based tool that were not a focus in the rating decision. Teams tended to simply document them as part of the COS discussion.

- No statistically significant differences were observed in categorizing skills into outcome areas on the documentation for EI and ECSE programs or entry and exit meetings.

- A number of teams were not documenting skills as evidence on the form, but those that did generally described skills relevant to the outcome area.

Exhibit 5.42 Documentation Errors About Skills Related to Child Outcomes on COS Forms

	Entire Sample (*n* = 113) *n* (%)	Subset with Evidence Including Skills (*n* = 94) *n* (%)
No major errors categorizing skills related to outcomes	91 (81)	91 (97)
Number of forms with Major Errors	3 (3)	3 (3)
Number of forms with no skills listed in evidence	19 (17)	Not Included

Summary, Implications, and Limitations for Study 3

The data from the team decision-making study, which involved extensive coding of 131 actual COS team meetings (63 for EI and 50 for ECSE) as well as a review of the associated documentation, provided a rich picture of how teams of providers discussed children's skills, behaviors, and functioning to derive the COS ratings. The data generally showed wide variation in implementation. Parents' or family members' participation in the COS decision-making meeting did not result in a more systematic approach to the COS process or consistent explanations with the family; instead, implementation continued to be variable. COS meetings varied considerably in length and depth of discussion and in comprehensiveness of the description provided about the child's skills. A small group of COS meetings involved a fairly quick affirmation of the rating with the parent, documenting an assessment, and little discussion or rationale on the videos for the COS ratings made.

However, we also obtained a variety of evidence that in somewhat longer meetings where more information was made available, the COS process can be done well and yield accurate data. A number of the findings support this general conclusion.

- The COS ratings for the videos coded spanned the full range of COS ratings, and correlations between ratings on the three outcomes were in the moderate to high range.
- While the length of time spent to derive the COS ratings was less than 5 minutes for a small percentage of videos, the majority of them ranged between 5 and 17 minutes. Importantly, coder ratings of the overall quality of the team process using a 5-point scale indicated that a majority of overall quality ratings were in the mid to high range. Not surprisingly, teams that spent longer on the COS process were rated by coders as having a higher overall quality COS process.
- Having a quality team process appears to support a quality COS process, as indicated by coders ratings about how well teams demonstrated use of COS guidance.
- Teams generally demonstrated skills that are important for a quality COS process.
  - o Teams only rarely made errors in categorizing skills and behavior into the correct outcome.
  - o Most videos showed teams having good to moderate breadth of discussion of the three child outcomes.

- o Most teams referenced functional skills to some degree in discussing the three outcomes, but only in about half or a little over half of the videos did most of the team's reference functional, rather than discrete, skills (i.e., 58%, 47%, and 61% for positive social relationships, knowledge and skills, and action to meet needs, respectively).
- o For all three outcomes, most teams discussed skills and behaviors across settings.
- Most teams (80%) made use of the COS decision tree, a resource developed by the ECO Center to assist teams in accurately distinguishing between the different rating points on the COS 7-point scale.
- In most videos (81%), teams referenced criteria language from the COS decision tree indicating appropriate age anchoring about skills and behaviors being considered in making ratings. Additionally, most videos (84%) had no age-anchoring errors.
- Problematic application of the rating criteria with regard to age anchoring was rare (i.e., only seen on 10%, 2%, and 0% of videos for positive social relationships, knowledge and skills, and action to meet needs, respectively).
- Misapplications of rating criteria other than for age anchoring occurred on 40% of the videos for at least one outcome. These errors often happened when ratings were based on progress rather than distance from age-expected skills or from overemphasis on one aspect of skills within an outcome.
- In very few videos (five, 4%) were there instances where one or more team members indicated that the ratings would make the program look good, and all involved inflating exit ratings.

Several limitations of the team decision-making study include the following.

- The sample of videos was small and not representative of the population of children served in EI and ECSE.
- The small sample also made it impossible for the research team to disaggregate data for analysis of various important predictor variables.
- It was clear to the coders that the videos often did not reflect all the knowledge that the team members bring to the COS rating process, as noted in the ways that team members who work together regularly may speak in a shorthand possibly because of the shared knowledge about the children they serve.
- Furthermore, for some videos of the COS rating process that occurred at the end of and IFSP or IEP meeting, there were indications on the video that some of the critical information that team members used to make ratings and reach consensus had been discussed and shared in the earlier part of the meeting that was not recorded.
- As with any study with video-recording, there are unknown effects from individuals knowing that they were being videotaped.

Section 6 — Study 4: Examining Validity Through Extant Data – Design, Methods, Key Findings

Research Questions for Study 4

In this study, the research team examined the validity of COS ratings by analyzing state-level data from several states using the COS process for OSEP reporting. To test the validity of the COS data across multiple states, the patterns in the state data were compared with expectations for valid data. Data are reported from two samples of states. The first sample comprised states that agreed to share data with the research team for the study, whereas the second sample was composed of all states using the COS for child outcomes reporting. The data were analyzed to address research questions about the extent to which existing state data support a subset of the validity claims. The claims examined with extant data were as follows.

- Claim C: There is variability in children's functioning in the three outcome areas, and that variability is reflected in the COS ratings.
- Claim F: Functioning in one outcome area is related to functioning in another outcome area.
- Claim H: COS ratings will be related to the type and severity of the child's disability.
- Claim K: COS rating distributions at entry will be related to the disability-related characteristics of the population served by states.
- Claim L: Similar populations of children enter programs each year, so functional levels (COS ratings) should remain constant without intervening factors (e.g., new eligibility criteria, rigorous quality assurance, or improvement process implemented).
- Claim M: Functioning, as reflected in the COS rating, in an outcome area at time 1 is related to functioning in that outcome area at a later point in time.
- Claim N: The rating structure of the COS is sensitive to both improvements in and maintenance of developmental trajectories that occur in effective programs (i.e., COS ratings differentiate effective from ineffective programs).
- Claim O: Data produced by COS are sufficiently precise to allow states to track overall status of their EI or ECSE system with the summary statements and monitor change toward targets on those summary statements.

Sample for Study 4

A set of criteria for minimally acceptable child outcomes data quality had been developed and used by the research team to develop national estimates of outcomes for children who had received early intervention or early childhood special education. The criteria used for data quality were an acceptable level of missing data and no out-of-range values for the progress categories (see ECTA [2014] for details on the criteria). This work was done in the research team members' capacity of national technical assistance providers to state agencies in outcomes measurement.

The same criteria were used to identify states with acceptable levels of data quality to invite to participate in the extant data study. Thirty-seven EI and ECSE coordinators or data managers were contacted for participation. Of these, 18 (49%) sent at least 1 year of data; 9 were EI programs, and 9 were ECSE programs. The number of children included in the data sets provided by states ranged from 809 to 20,390. For all outcomes, a mean of 98% of children had made some progress while in early intervention programs and a mean of 99% of children in preschool special education programs had made some progress (i.e., these children were not in progress category "a"). Exhibit 6.1 shows the mode of the COS rating distribution complied across the data sets provided by states. In general, children across outcomes and programs entered the program functioning below age expectations (COS rating of 5 or below) and exited functioning at or above age expectations (6 or above). Children in EI entered with higher ratings in positive social emotional skills, whereas children in ECSE entered with higher ratings in taking actions to meet needs. Exhibit 6.2 shows the average entry ratings for EI, ECSE and overall.

Exhibit 6.1   Mode of COS Ratings at Entry and Exit from State Data Sets for EI and ECSE Included in the Extant Data Study

	Positive Social Relationships		Knowledge and Skills		Actions to Meet Needs	
	Entry	**Exit**	**Entry**	**Exit**	**Entry**	**Exit**
EI	5	6	3	6	3	6
ECSE	3	6	3	6	5	7

Exhibit 6.2   Average COS Ratings at Entry for State EI and ECSE Programs in Extant Data Study

At Program Entry	*n*	Positive Social Relationships Mean (SD)	Knowledge and Skills Mean (SD)	Action to Meet Needs Mean (SD)
Overall mean (SD) across all EI state programs	30,625	4.6 (1.6) Range: 3.4 – 4.3	4.0 (1.5) Range: 3.3 – 4.3	4.2 (1.5) Range: 3.8 – 4.8
Overall mean (SD) across all ECSE state programs	61,108	3.9 (1.7) Range: 3.6 – 4.9	3.8 (1.6) Range: 3.3 – 4.5	4.2 (1.8) Range: 3.2 – 4.7
Overall mean (SD) across all extant data states (EI/ECSE combined)	91,678	4.1 (1.7) Range: 3.4 – 4.9	3.9 (1.6) Range: 3.3 – 4.5	4.2 (1.7) Range: 3.2 – 4.8

A limitation of the analyses of the state data from both samples was that little is known about the quality of implementation and the quality of data in these states. Although the states met the criteria for data quality for inclusion in the national analyses, these criteria are minimal and states meeting them could have problems with implementation and data quality.

Data Analysis for Study 4

Analysis of Data Sets Provided by States

The research team used the most recent year of data submitted by each state to examine support for validity claims. The most recent year of data provided by states ranged from data collected on children exiting the programs between summer 2009 and summer 2010 to those who exited between summer 2012 and summer 2013. As indicated below, for some validity claims multiple years of data from the same state were used.

Analysis of National Data

Federally reported child outcomes data for all states using the COS were analyzed for those claims that could be addressed with the level of data states report to the federal government. National child outcomes data are compiled annually by the Early Childhood Outcomes center and the Early Childhood Technical Assistance center as part of a review of states' Annual Performance Reports, and these data were used for these analyses.

Determining Evidence for Claims

Criteria were developed to reflect the expected patterns in state data that would support the validity of the COS. Data from the state data sets and national data were analyzed to determine whether each state's data met the criteria. The development of the criteria was informed by the expected properties of the measurement scale, populations of children with disabilities served by states, and findings from developmental research. For each analysis, the percentage of states that met established criteria was tallied. This approach was selected because very little was known about the quality of any individual state's data, and compiling data across states would have resulted in larger states having more influence on the findings than smaller states.

The first step in conducting each analysis was to set the criteria for validity. After the criteria were set, data from all states were reviewed and categorized as either

- **Yes,** the state met the criteria for the claim, or
- **No,** the state did not meet the criteria.

For each claim, outcome, and program (EI or ECSE) the percentage of states that met the criteria was computed. Each claim was tested across two programs and three outcomes. If more than 50% of states included in the analysis met the criteria for all outcomes and programs, the claim was considered supported. If more than 50% of states met the criteria for some outcomes for at least one of the programs, the evidence was considered mixed. If less than 50% of states met the criteria for all outcomes and both programs, the claim was not supported. In sum, claims were considered

- **Supported** if a majority of the states met the criteria across all outcomes and both programs.

- **Mixed** if the results were met for some, but not all, outcomes or for both programs.
- **Not supported** if the majority of states did not met the criteria across all outcomes and both programs.

Key Results for Study 4: Extant State Data Study

Claim C: There is variability in children's functioning in the three outcome areas, and that variability is reflected in the COS ratings

Children served in EI and ECSE enter and exit programs with a wide range of abilities and skills and also make varying degrees of progress between entry and exit. The 7-point rating scale on the COS was intended to reflect a range of abilities, from the lowest rating of 1 to the highest of 7.

We expected that that very few children served by EI and ECSE have severe impairments in any outcome and therefore that the percentage of children with ratings of 1 should be very low. Likewise, we expected that relatively few children requiring special education services would be functioning at a level comparable to same-age peers across all three of the functional outcomes. With regard to the COS ratings, this meant that percentages of children with ratings of 6 or 7 should be low at entry. Children entering with ratings of 6 or 7 on all three outcomes on the COS would usually be those who function effectively in everyday environments with assistive technology but who may not function as effectively without those assistive technology supports. Another example would be children with diagnosed conditions or physical/sensory impairments who enter the program functioning at age expectations and receive supportive services to keep their disability from affecting skill development and functioning in other areas.

Three subclaims were identified for Claim C. The subclaims addressed the range of scores at entry and exit and the skill level at entry.

Subclaim C1: Children being served in EI and ECSE will demonstrate a full range of scores on the COS at entry and exit

The criterion to meet subclaim C1 was that state data would include a full range of scores (1 to 7) on the COS at entry and exit. The rationale for this expectation was that a full range of functional skills and abilities should be present among children receiving early intervention or special education services in a state.

For both EI and ECSE, 100% of states demonstrated the full range of COS ratings in their data at both entry and exit. This subclaim was supported because a majority of the states met the criterion.

Subclaim C2: On each outcome, the majority of states will show that less than 10% of children being served in state EI and ECSE programs enter with ratings of 1 on the COS.

For the state data to be considered as supporting this subclaim, less than 10% of children in the state data had to be rated as a 1 on the COS rating scale. A rating of 1 is defined as: "Child

does not yet show functioning expected of a child his or her age in any situation. Child's functioning does not yet include immediate foundational skills upon which to build age-appropriate functioning. Child functioning reflects skills that developmentally come before immediate foundational skills." The rationale for this criterion was that very few children served by EI and ECSE have severe impairments in an outcome area.

For EI, all states met the criteria for positive social relationships, and 6 out of 9 states (67%) met the criteria for knowledge and skills and takes action to meet needs. For ECSE, 7 out of 9 states (78%) met the criteria for positive social relationships and knowledge and skills, and 8 out of 9 states (89%) met the criteria for takes actions to meet needs. See Exhibit 6.2. This subclaim was supported because a majority of the states met the criteria across all outcomes for both programs.

Exhibits 6.2   Percentage (Number) of States in EI and ECSE with Fewer than 10% of Children Rated a 1 at Program Entry (*n* = 9 EI and 9 ECSE state programs)

Program	Positive Social Relationships		Knowledge and Skills		Takes Action to Meet Needs	
	Number	Percent	Number	Percent	Number	Percent
EI	9	100%	6	67%	6	67%
ECSE	7	78%	7	78%	8	89%

Subclaim C4: Very few children will enter programs functioning at a level comparable to same-age peers across all three child outcomes

For the state data to be consistent with subclaim C3, fewer than 15% of children had to enter EI or ECSE functioning at age expectations across all three outcomes. The rationale for this subclaim was that very few children found to be eligible for early intervention or preschool special education services are likely to be functioning at a level comparable to same-age peers across all three of the functional outcomes when they enter the program.

Data from one state could not be used in this analysis, bringing the total number of states for EI and ECSE to eight. For EI, all states had fewer than 15% of children entering at age expectations on all three outcomes. For ECSE, six of eight states had fewer than 15% of children entering at age expectations on all three outcomes (Exhibit 6.4). A majority of the states for both programs met the criteria.

Exhibit 6.4   Percentage of EI and ECSE States with Fewer than 15% of the Children Entering the Program at or Above Age Expectations on All Three Outcomes

Program	Percent (Number) of States
EI	100% (8 of 8)
ECSE	75% (6 of 8)

Claim F: Functioning in an outcome area is related to functioning in the other outcome areas

The criterion for this claim was moderate to strong correlations (.5 or above) between COS ratings across outcomes at entry and at exit. The rationale for this expectation was that development in early childhood is integrated across developmental areas such that development in one outcome area for many children progresses in tandem with development in other areas. Claim F had one subclaim that was tested at entry and exit.

Subclaim F1: COS ratings will show moderately strong correlations (.5 or stronger) across outcomes at entry and at exit

This claim was supported because for both EI and ECSE, all 9 EI states and all 9 ECSE states met the criterion for the claim across all outcomes at entry and exit. Across the three outcome pairs, correlation ranges at entr were .70-.83 for ECSE and .55-.80 for EI. At exit, correlation ranges were .70-.87 for ECSE and .74-.88 for EI.

Claim H: COS ratings will be related to the type and severity of the child's disability

Children participating in ECSE are assigned a primary disability based on a set of federally defined categories. EI programs do not use disability categories, so they were not included in this analysis. Because only two ECSE states provided COS rating by primary disability, this claim was tested just in two states. The subclaims addressed the relationship for two primary disability categories: autism and speech or language impairment.

Subclaim H2: Children with a primary disability of autism will have lower ratings in positive social emotional skills relative to the other two outcome areas and relative to children with other disabilities

To meet the criteria for subclaim H2, the state data would show that children with a primary disability of autism (1) received lower ratings in social emotional skills than children with other disabilities and (2) were rated lower in social emotional skills than in knowledge and skills or actions to meet needs. The ratings of children with autism were compared with those for a comparison group composed of all other primary disabilities.

- In state 1, 99% of children with a primary disability of autism were rated below age expectations in positive social emotional skills compared with 75% of children with other primary disabilities.
- In state 2, 99% of children with a primary disability of autism were rated below age expectations in positive social emotional skills compared with 90% of children with other primary disabilities.
- In state 1, 99% of children with a primary disability of autism were rated below age expectations in positive social emotional skills compared with 92% rated below age expectations in knowledge and skills and 92% rated below age expectations in actions to meet needs.

- In state 2, 99% of children with a primary disability of autism were rated below age expectations in positive social emotional skills compared with 99% rated below age expectations in knowledge and skills and 96% rated below age expectations in actions to meet needs.

Evidence from both states supported the subclaim that children with a primary disability of autism would have lower ratings in positive social relationships. Evidence for ratings in positive social relationships being lower than those for the other two areas differed across the states. Data from state 1 indicated that children with a primary disability of autism were rated lower in positive social relationships than they were in the other two outcome areas. The data from state 2 did not show that children with a primary disability of autism were rated lower in positive social relationships compared to ratings for the same children in knowledge and skills and actions to meet needs.

Subclaim H3: Children in ECSE with a primary disability of speech or language impairment will have significantly higher COS ratings on each outcome compared with children with other primary disabilities

For this subclaim to be supported, children with a primary disability of speech language impairment had to be rated higher for each outcome compared with children with all other primary disabilities. This subclaim was examined in the same two states used for subclaim H3. Exhibit 6.4 shows the comparison of children with a primary disability of speech or language impairment to children with other primary disabilities.

Exhibit 6.4 **Percentage of Children with a Speech or Language Impairment Rated as Below Age Expectation for Each Outcome Compared with Children with Other Primary Disabilities in Two States**

	State and Disability Group			
	State 1		State 2	
Outcome Area	**Speech or Language Impairment**	**Other Disabilities**	**Speech or Language Impairment**	**Other Disabilities**
Number of children	2086	4110	Range across outcomes: 2256-2290	Range across outcomes: 2903-2916
Positive social relationships	51	90***	84	95***
Knowledge and skills	54	89***	85	94***
Action to meet needs	33	83***	76	90***

*p < .05, ** p < .01, *** p < .001, **** p < .0001.

In both states, fewer children with a primary disability of speech or language impairment were rated below age expectations compared with children with other disabilities. This is consisten with what was expected. Both subclaims, which addressed differences in ratings across different disabilities, were supported and therefore Claim H was supported.

Claim K: COS rating distributions at entry will be related to the disability-related characteristics of the population served by states

There are state-specific variations in requirements for eligibility for Part C and Part B Preschool program participation. Some states require the child's disability or delay to be more severe than others in order to be eligible to receive services. State Part C programs have latitude in defining both the range of disabilities and conditions that qualify children for services and the threshold of impact on the child's functioning at entry in order to qualify for services (e.g., extent of documented developmental delay required and number of domains in which delay is evident). Part B Preschool uses the same set of primary disability categories nationally, but states have discretion in setting the threshold of impact on the child's functioning at entry in order to qualify for services.

It was expected that states with broader eligibility guidelines would serve a higher percentage of children in the state than states with more narrow, stringent requirements. States using the COS identify ratings for all children as they enter Part C services. If these ratings capture children's level of functioning effectively, then one would expect to see that states with more stringent eligibility guidelines serving a smaller percentage of children (i.e., providing services only to children who have more severe impairments to their functioning) would have lower COS ratings on average than states serving a higher percentage of children with a broader range of severity in their impairments. This claim assumed that ChildFind activities would be comparable from one state to another and that states with low percentages served were serving fewer children because of stricter eligibility requirements, not because of limited outreach to find families and children who need services.

Subclaim K1: National analysis comparing the percentage of young children served by EI or ECSE to the percentage of children exiting EI or ECSE at age expectations for states using the COS in 2012-13

The expectation was that national data comparing states using the COS process would show a positive relationship between the percentage of children in the state served by the EI or ECSE program and the percentage of children who exited the program functioning at age expectations on each outcome.

Exhibits 6.5 and 6.6 show the relationships between the percentage of children that the COS identified as exiting at an age-expected levels on each outcome in 2012–13 and the percentage of children that the state EI or ECSE program reported serving in the state.

Exhibit 6.5   **Children Exiting EI at Age-Expected Levels by Outcome and Children the ECSE Program Reported Serving in the State (2012–13) (*n* = 37 States)**
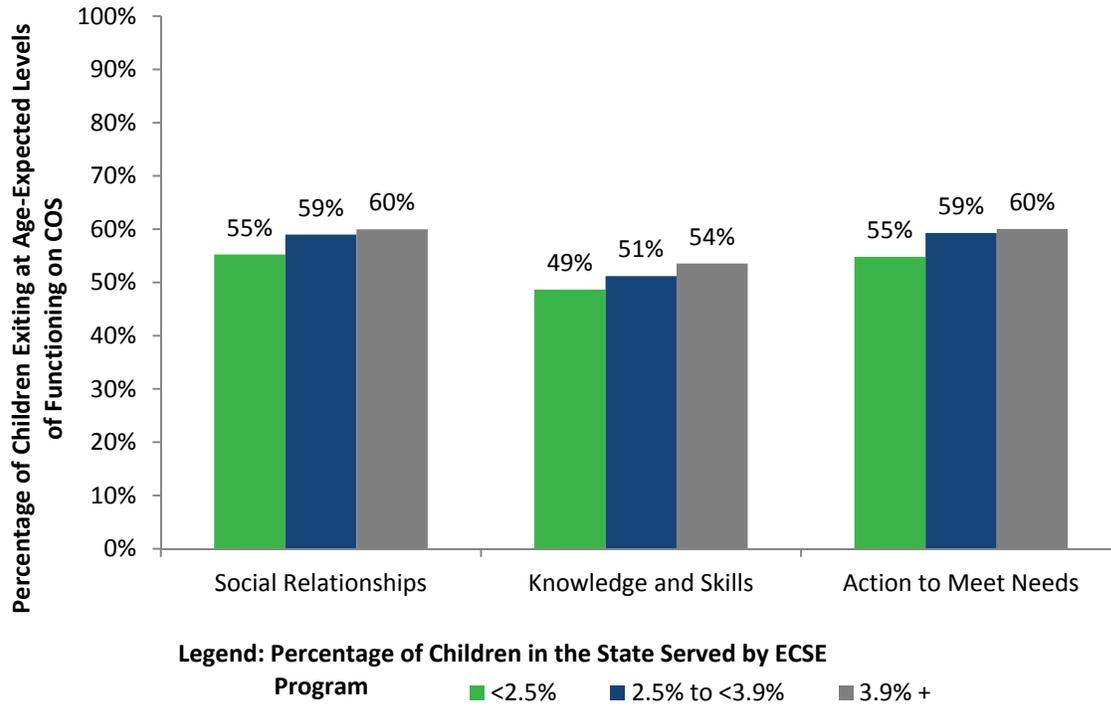
Note: Social relationships *n* = 15, knowledge and skills *n* = 13, and action to meet needs *n* = 9.

Exhibit 6.6   Children Exiting ECSE at Age-Expected Levels by Outcome and Children the ECSE Program Reported Serving in the State (2012–13) (*n* = 29 states)
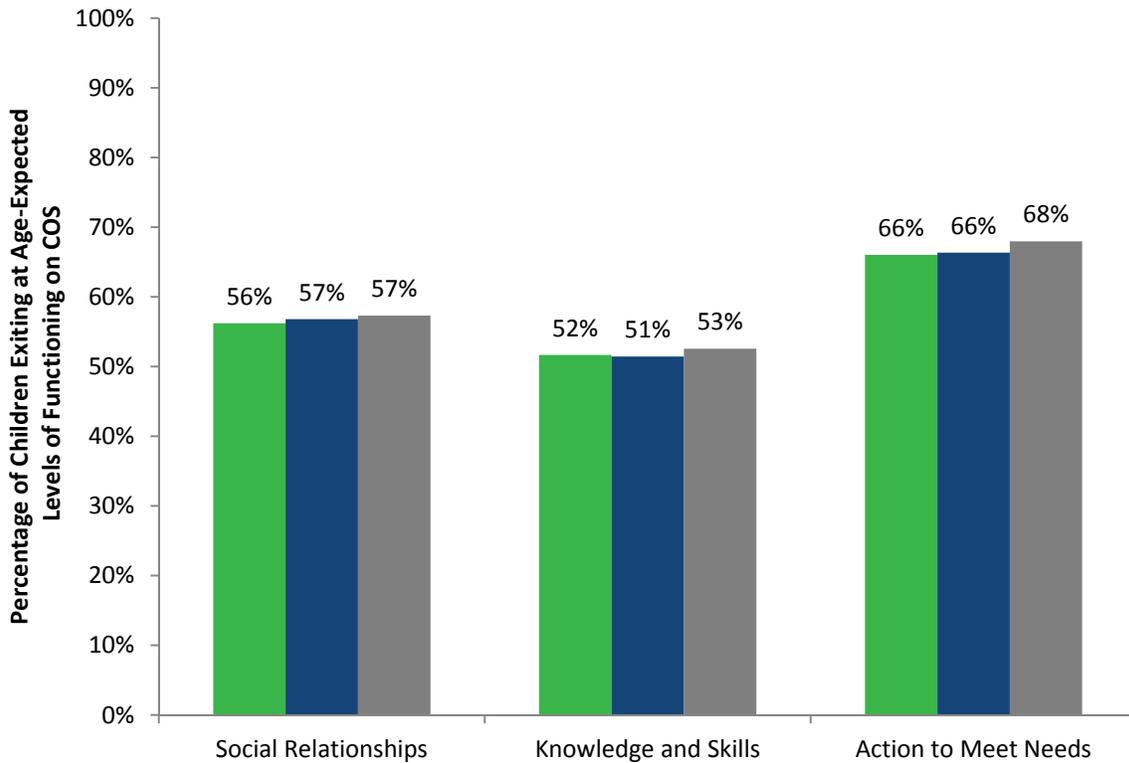
Legend: Percentage of Children in State Served by ECSE Program
■ <5.7% (n = 11)   ■ 5.7-7.5% (n = 11)   ■ >7.5% (n = 7)

- The national data for 37 EI programs demonstrated the expected stair-step pattern of increasing percentages in states serving a higher percentage of children in the EI program (i.e., if a state serves more children, then the populations served should contain more higher functioning children). Given the small number of states divided across three categories, the mean differences in percentages did not reach the threshold for statistical significance. However, the pattern of relationships was consistent with the criterion (despite limitations of relationships between strictness of eligibility requirements and percent served by the state).

- The national data for 29 ECSE programs showed less variation across the three groups. States with the highest percentage served did have the highest percentage of children at age-expected levels on each outcome. However, there was less distinction between states with lower levels of percentage of children served. Weaker data for ECSE were expected given less diversity in state-specific discretion about children served in ECSE than EI programs. However, in ECSE, the expected stair-step pattern was not supported.

Claim L: Similar populations of children enter programs each year, so functional levels (COS ratings) should remain constant without intervening factors (e.g., new eligibility criteria, rigorous, quality improvement efforts)

This claim was considered supported if there was no more than a 3-point change in the percentage of children rated at each of the COS ratings at entry from year to year for 80% (17 of the 21) comparisons in a state. Overall, it was considered supported if 80% of the states met this criterion.

Exhibit 6.7 shows the largest percentage change from year to year in entry frequencies for each point on the rating scale for three EI and two ECSE state programs. The year-to-year change was calculated by subtracting the percentage of children who received that rating in one year from the percentage who received it in the later year (e.g., 17% in 2010–11 minus 16% in 2009–10 gives a change of 1%).

There were no year-to-year changes of greater than 3 points for any of the ECSE states. For the EI programs, there were changes of greater than 3 points between 2008–09 and 2009–10 in state 2 for 38% of the comparisons. State 2 is a small state with COS ratings on approximately 500 children, which means that the cell sizes for each of the individual ratings was relatively small and therefore more subject to year-to-year fluctuations. For the larger states, the year-to-year distributions were incredibly stable across all the rating points. Eighty percent of the states (4 of 5) meet the criteria for this claim.

Exhibit 6.7   Largest Year-to-Year Change (Percentage Minus Percentage) for Each of the Rating Points for Each Outcome for EI and ECSE

	COS Ratings						
	1	2	3	4	5	6	7
EI: State 1 (Years: 2008 – 2010; *n* = 2,721 to 3,556)							
Positive social relationships	1	0	2	1	1	2	1
Knowledge and skills	1	0	1	1	1	2	1
Action to meet needs	1	0	2	0	1	0	1
EI: State 2 (Years: 2008-9 to 2010-11, *n* = 431 to 606)							
Positive social relationships	2	1	1	2	**7**	**4**	**4**
Knowledge and skills	1	1	1	3	**5**	2	1
Action to meet needs	3	**4**	2	**4**	2	**4**	**6**
EI: State 3 (Years: 2008-09 to 2009-2010, *n* = 4,141 to 4,242)							
Positive social relationships	0	1	2	0	1	1	0
Knowledge and skills	1	1	0	0	2	1	0
Action to meet needs	1	0	3	1	0	1	1
ECSE: State 1 (Years: 2008-09 to 2010-11, *n* = 6,196 to 7,256)							
Positive social relationships	0	2	0	0	2	1	1
Knowledge and skills	0	1	1	1	1	0	0
Action to meet needs	1	1	2	1	2	1	1
ECSE: State 2 (Years: 2008-09 to 2010-11, *n* = 5,906 to 10,124)							
Positive social relationships	1	1	2	2	1	2	1
Knowledge and skills	1	1	2	1	2	2	1
Action to meet needs	2	2	3	2	1	3	2

Claim M: Functioning in an outcome area at entry is related to functioning in that area at exit

Functioning in an outcome area at entry, as reflected in the COS rating, was expected to be related to functioning in that area at exit.

Subclaim M1: COS ratings in the same outcome area at time 1 and time 2 will show moderate correlations (.5 or greater)

The criterion for this subclaim was moderate to strong correlations between COS ratings at entry and exit (.5 or above). The rationale for this expectation was that many children display stability in their development over time (e.g., children who enter the program with lower levels of functioning continue to have difficulties in that outcome area relative to children who enter with higher levels of functioning).

For ECSE, all states met the criteria for the subclaim across all outcomes. For EI, six out of nine states (67%) met the criteria for positive social relationships, and three out of nine states (33%) met the criteria for knowledge and skills and takes action to meet needs.

As shown in Exhibit 6.8, the evidence for this subclaim was mixed because it was met by all states for ECSE but by fewer than 80% for EI.

Exhibit 6.8 **Percentage and Number of EI and ECSE States with COS Ratings with Correlation Coefficients of .5 or Above Between Entry and Exit for Each of the Three Outcomes**

Program	Positive Social Relationships	Knowledge and Skills	Takes Action to Meet Needs
EI	67% (6 of 9)	33% (3 of 9)	33% (3 of 9)
ECSE	100% (9 of 9)	100% (9 of 9)	100% (9 of 9)

Subclaim M2: Twenty percent or less of children will show a change of 4 or more rating points between time 1 and time 2

Programs are designed to help children achieve changes in overall functioning. Extremely large changes, however, are suspect and suggest the data may not be valid. Most children in the state should maintain or improve their status relative to same-age peers, but few should show large increases between entry and exit. Large changes in status relative to same-age peers between entry and exit from the program are possible but rare. For this subclaim to be considered supported, 20% or less of children should have more than or equal to a 4-point increase in their COS rating between entry and exit.

Nine EI and nine ECSE states were included in the analysis. For both EI and ECSE, 100% of states had 20% or less children showing increases of more than 4 points between entry and exit.

Claim N: The rating structure of the COS is sensitive to both improvements in and maintenance of developmental trajectories that occur in effective programs (i.e., COS ratings differentiate effective from ineffective programs)

Subclaim N1: COS ratings can detect a portion of children who change developmental trajectories between time 1 and time 2

National data also show that states were able to report on the percentage of children statewide who entered below age expectations and positively changed their developmental trajectory before exiting the program. All states using the COS provided this information to OSEP as part of annual reports. Exhibits 6.9 and 6.10 show national data, with the distribution of statewide percentages for Summary Statement 1 from all states using the COS process. These exhibits show percentages for EI and ECSE on the action to meet needs outcome as an example. Exhibit 6.11 summarizes the range of values observed across states on all the outcomes. State EI programs show variance in the percentage of children that make greater than expected growth.

Exhibit 6.9   Distribution of Summary Statement 1 Percentages Across State EI Programs for Action to Meet Needs Outcome (2012–13)
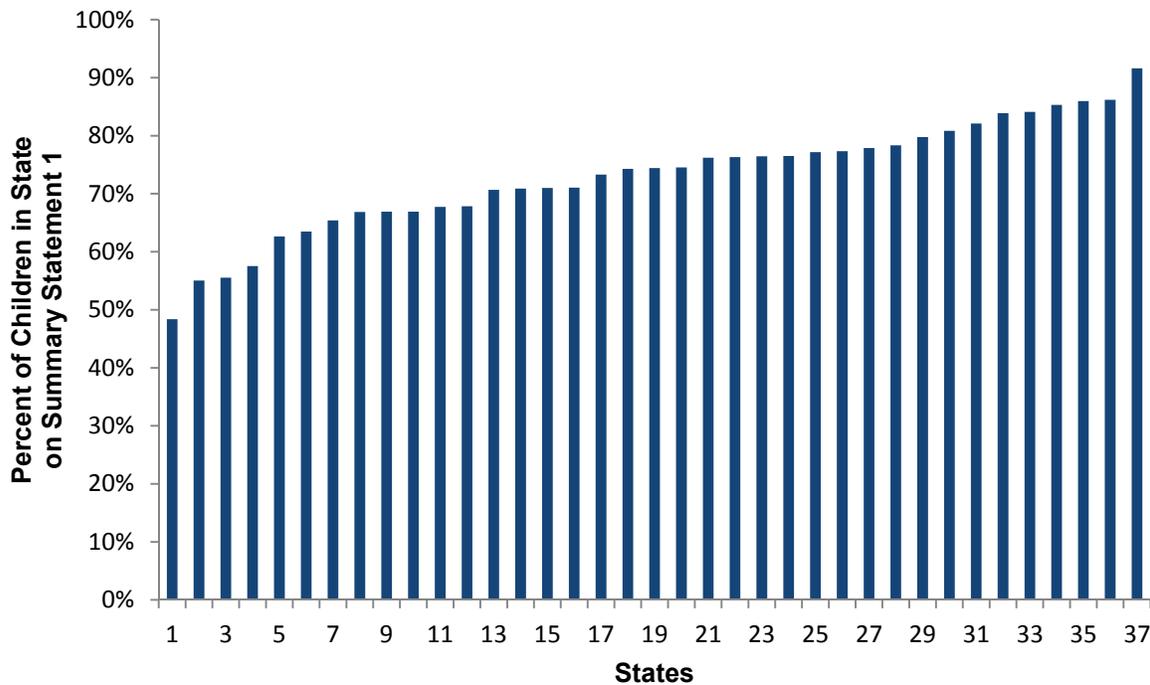
Exhibit 6.10 **Distribution of Summary Statement 1 Percentages Across State ECSE for Action to Meet Needs Outcome (2012–13)**
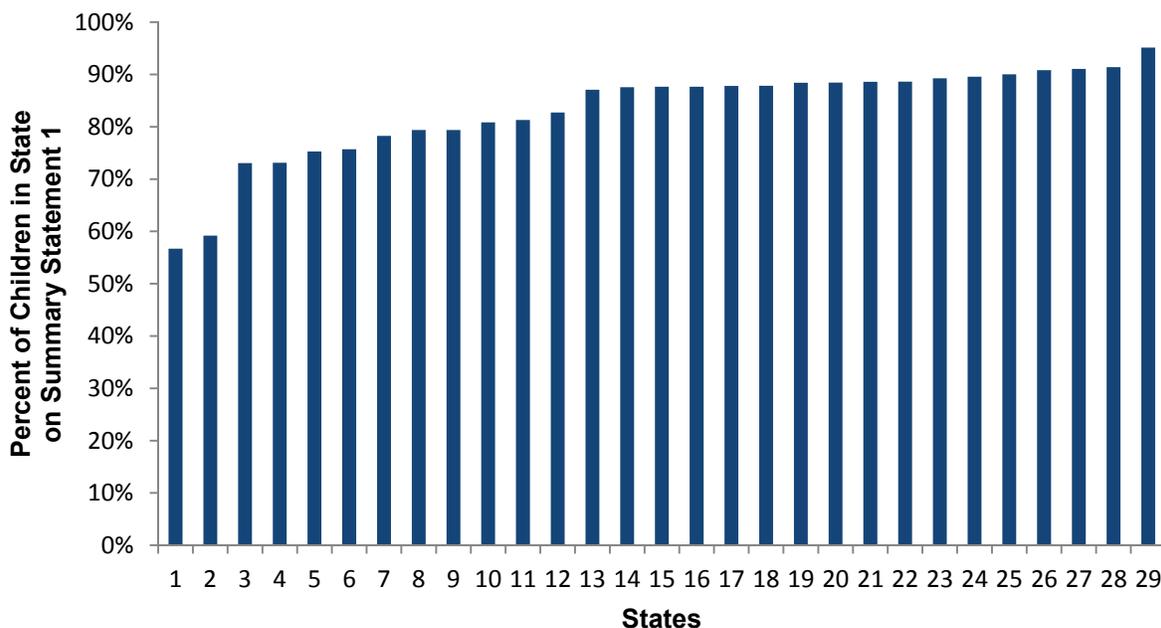
Exhibit 6.11 **Range of Summary Statement 1 Percentages Across State EI and ECSE Programs by Outcome (2012-2013)**

	Range of Percentages on Summary Statement 1 (Greater Than Expected Growth Developmental Trajectory Changes)		
	Positive Social Relationships	**Knowledge and Skills**	**Actions to Meet Needs**
EI programs (*n* = 37 states)	24–91	37–95	48–92
ECSE programs (*n* = 29 states)	46–96	64–98	57–95

Subclaim N2: Programs will show differences in the percentage of children making greater than expected growth

This subclaim was supported if programs within states showed variance in the percentage of children making greater than expected growth. To examine this claim, the range of Summary Statement 1 (greater than expected growth) values was examined across local programs. Programs had to have at least 30 children exiting to be included in the analysis. Seven EI state programs and eight ECSE programs provided data on local programs. As Exhibit 6.12 shows, as expected, there was variance across local programs for all states for each outcome in EI and ECSE.

Exhibit 6.12   Range in the Percentage of Children Making Greater Than Expected Growth Across Local Programs for EI and ECSE

State	Number of Local Programs Considered (*n* for each program >30)	Positive Social Emotional Skills Summary Statement 1	Knowledge and Skills Summary Statement 1	Actions to Meet Needs Summary Statement 1
EI				
1	2	78–89	83–87	76–90
2	23	52–97	56–98	63–98
3	17	46–80	40–78	55–86
4	19	21–63	18–56	26–63
5	17	94–100	61–96	56–92
6	20	56–88	59–90	59–94
7	7	66–94	68–94	73–93
ECSE				
1	10	70–97	71–96	71–99
2	102	0–100	0–100	0–100
3	16	72–100	74–100	79–98
4	42	55–93	49–95	55–97
5	2	83–86	83–86	85–87
6	41	76–100	72–100	69–100
7	71	21–100	19–98	18–98
8	22	79–100	75–100	78–100

Claim O: Data produced by COS are sufficiently precise to allow states to track overall status of their EI or ECSE system with the summary statements and monitor change toward targets on those summary statements

Subclaim O1: States will show stability in their summary statements from year to year. Few states will show patterns with erratic shifts in summary statements from year to year.

We would expect state summary statement values to remain stable across years because of the large number of children used to compute the estimate. We would expect changes only as a result of an effective statewide intervention, and these changes would be gradual. Few states would show patterns with erratic shifts in summary statements from year to year.

National summary statements submitted for federal reporting by state EI and ECSE programs were used to test this subclaim. Only data for states using the COS for accountability between 2009–10[32] and 2012–13 and meeting minimum quality criteria to be appropriate for the national analysis were analyzed. This resulted in use of data from 20 state EI and 18 state ECSE programs.

Few states made big shifts in Summary Statement 1 percentages (Exhibits 6.13–6.15). Most showed fairly stable or slightly incremental trajectories. For state EI programs with Summary Statement 1, this criterion was supported.

Exhibit 6.13   Pattern of Greater Than Expected Growth (Summary Statement 1) Percentages for Positive Social Relationships Across Years Among State EI Programs Using COS Process (2009–13) (*n* = 20 states)

---
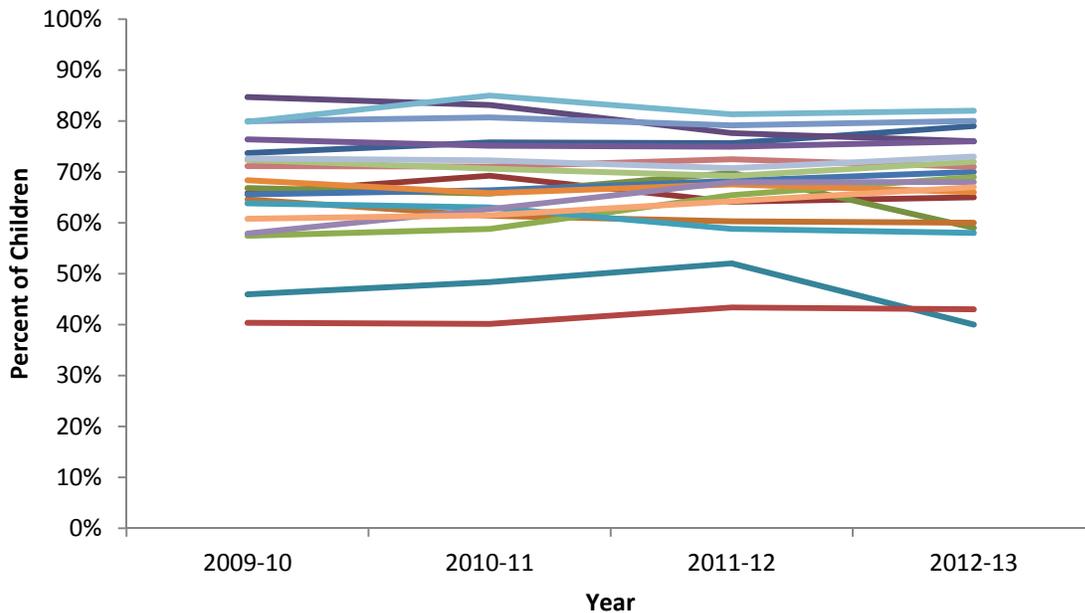
[32]   Most states had data from 2008–09, but sample sizes often shifted during those early years as states expanded outcomes data collection statewide and addressed major training issues influencing data quality. So data shown reflect 2009–2013.

**Exhibit 6.14   Pattern of Greater Than Expected Growth (Summary Statement 1)
Percentages for Knowledge and Skills Across Years Among State EI
Programs Using COS Process (2009–13) (*n* = 20 states)**
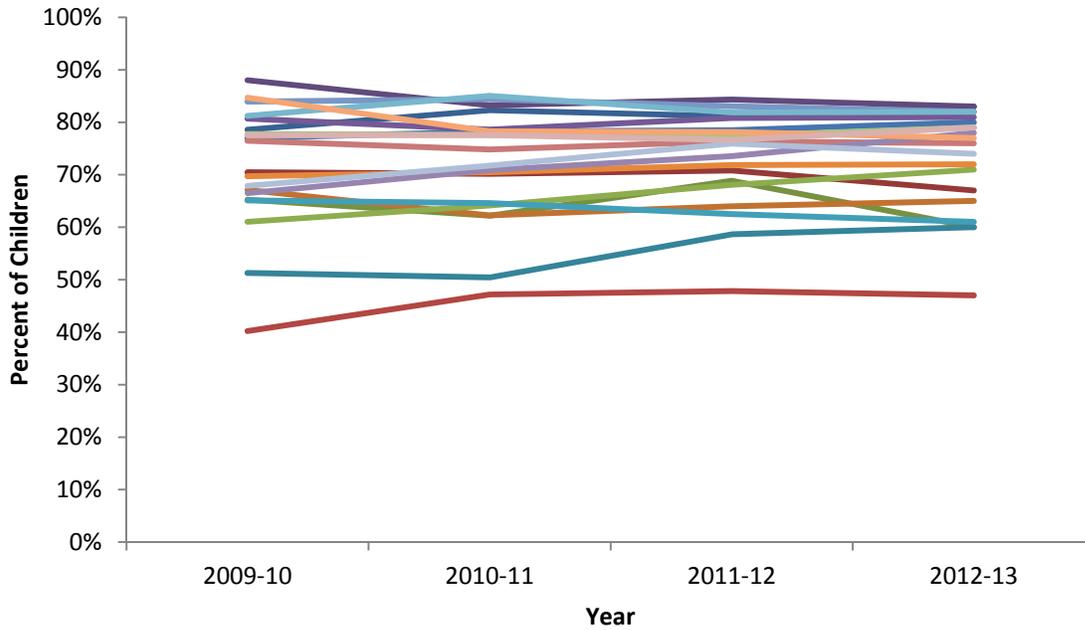
**Exhibit 6.15   Pattern of Greater Than Expected Growth (Summary Statement 1)
Percentages for Action to Meet Needs Across Years Among State EI
Programs Using COS Process (2009–13) (*n* = 20 states)**

182
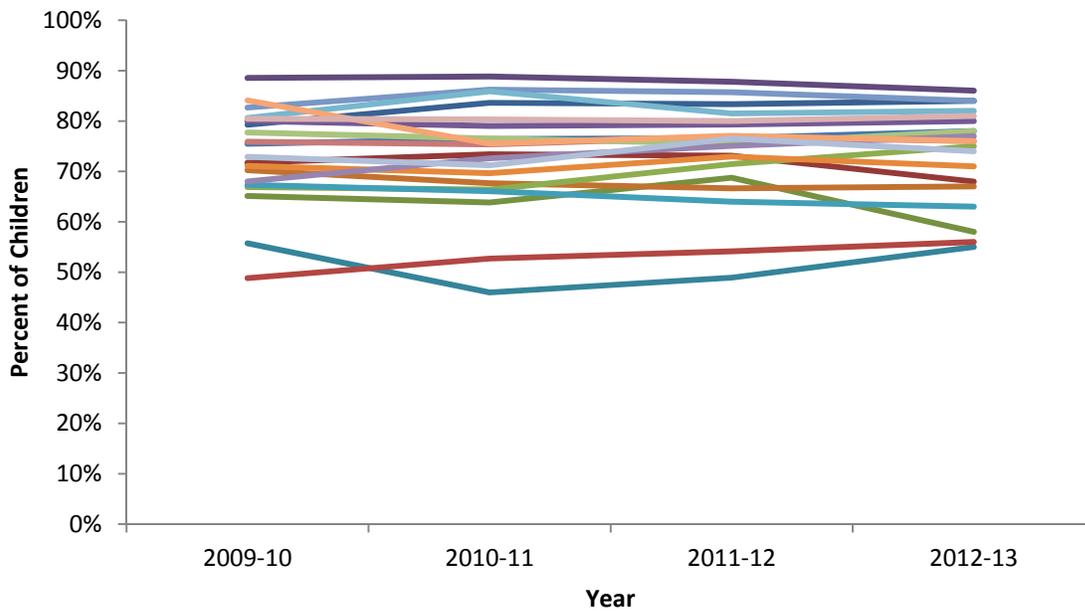
For state ECSE programs, few states made big shifts in Summary Statement 1 percentages (Exhibits 6.16 – 6.18). Again, most showed fairly stable or slightly incremental trajectories. For state ECSE programs with Summary Statement 1, Subclaim O1 was supported.

Exhibit 6.16   Pattern of Greater Than Expected Growth (Summary Statement 1) Percentages for Positive Social Relationships Across Years Among State ECSE Programs Using COS Process (2009–13) (*n* = 18 states)
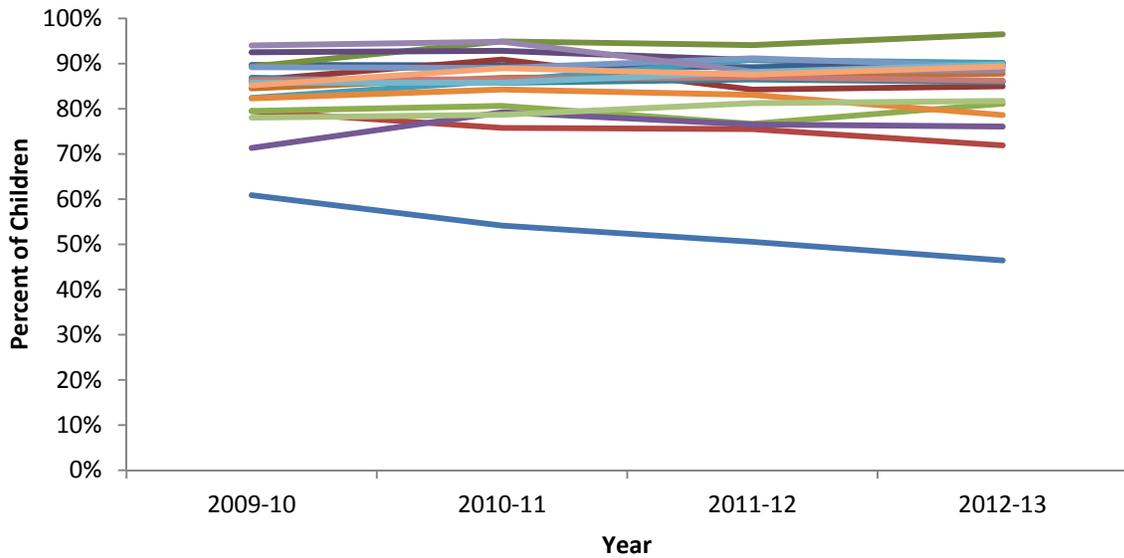
Exhibit 6.17   Pattern of Greater Than Expected Growth (Summary Statement 1)
Percentages for Knowledge and Skills Across Years Among State ECSE
Programs Using COS Process (2009–13) (*n* = 18 states)
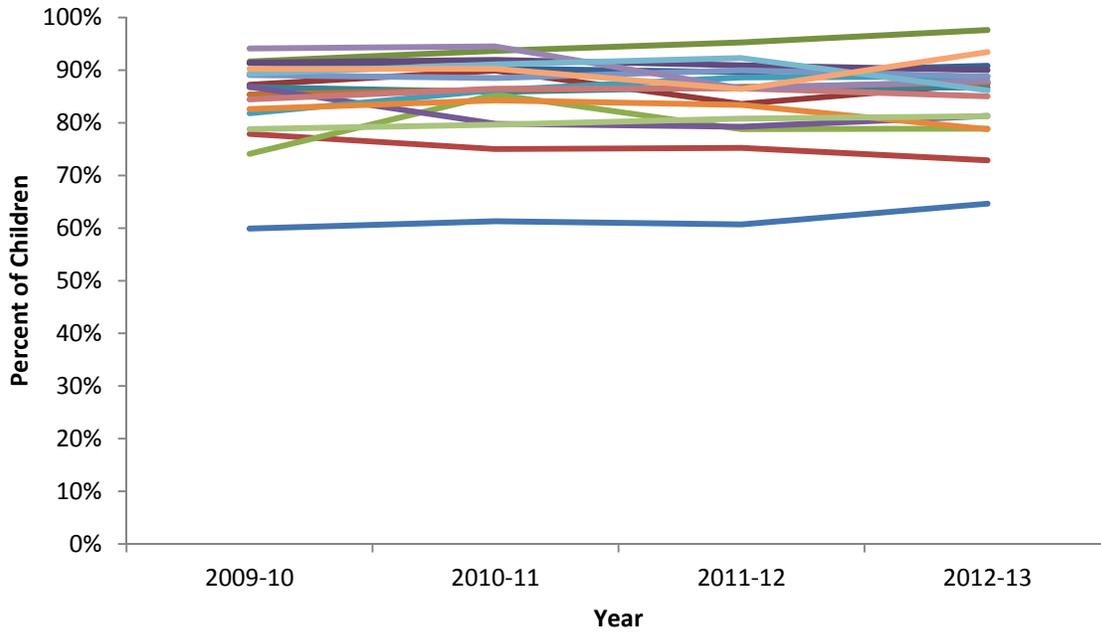
Exhibit 6.18   Pattern of Greater Than Expected Growth (Summary Statement 1)
Percentages for Action to Meet Needs Across Years Among State ECSE
Programs Using COS Process (2009–13) (*n* = 18 states)
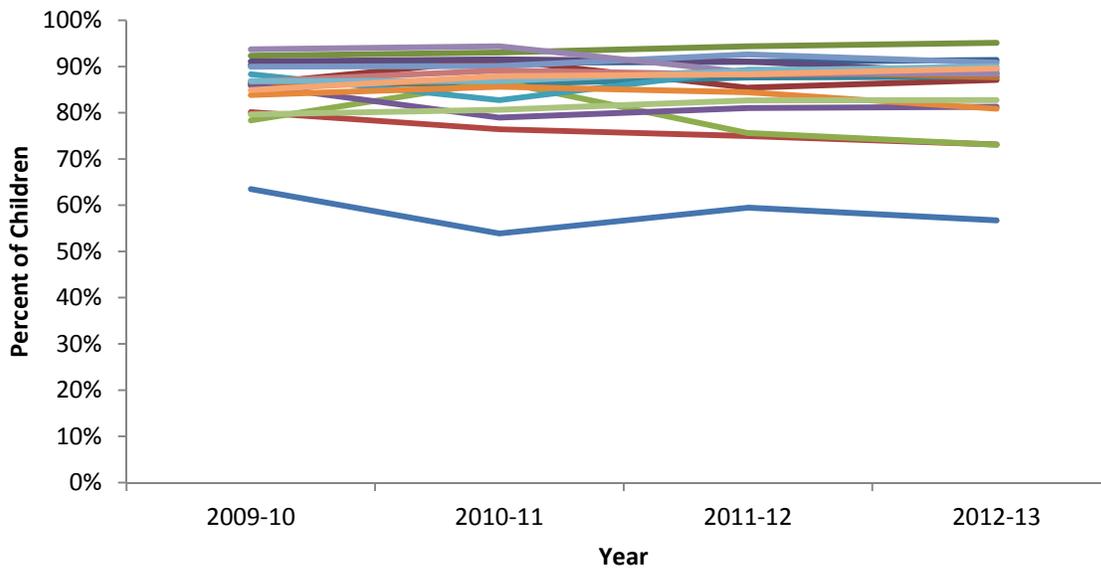
For state EI programs, few states had big shifts in Summary Statement 2 percentages (Exhibits 6.19 – 6.21). Most states showed fairly stable or slightly incremental trajectories. For state EI programs with Summary Statement 2, Subclaim O1 was supported.

Exhibit 6.19   Pattern of Exiting at Age Expectations (Summary Statement 2) Percentages for Positive Social Relationships Across Years Among State EI Programs Using COS Process (2009–13) (*n* = 20 states)
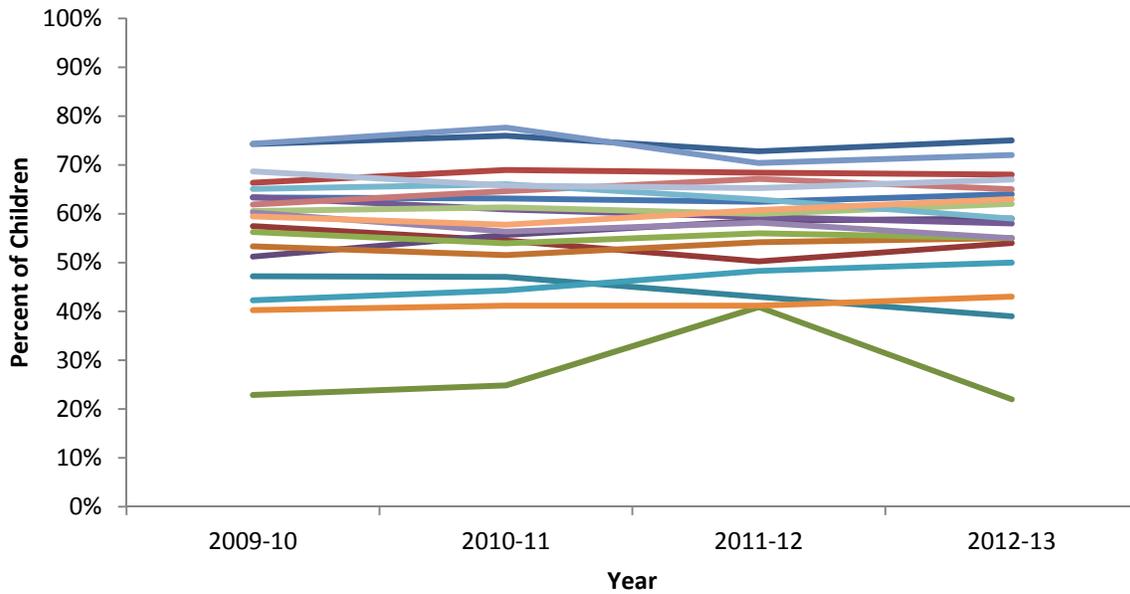
Exhibit 6.20   **Pattern of Exiting at Age Expectations (Summary Statement 2) Percentages for Knowledge and Skills Across Years Among State EI Programs Using COS Process (2009–13) (*n* = 20 states)**
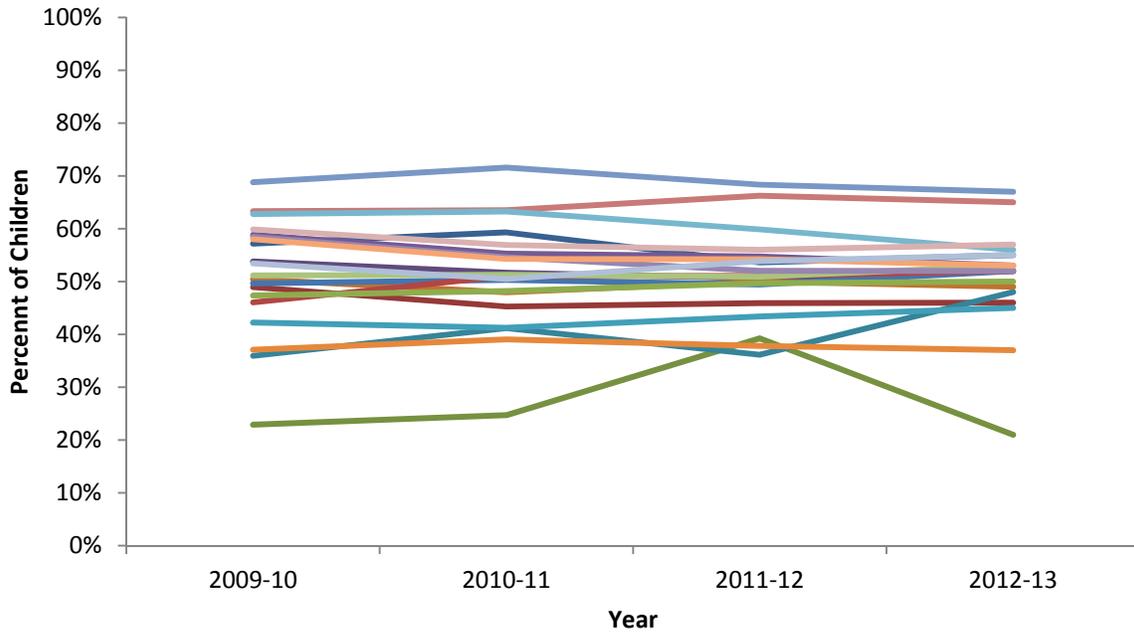
Exhibit 6.21   **Pattern of Exiting at Age Expectations (Summary Statement 2) Percentages for Action to Meet Needs Across Years Among State EI Programs Using COS Process (2009–13) (*n* = 20 states)**
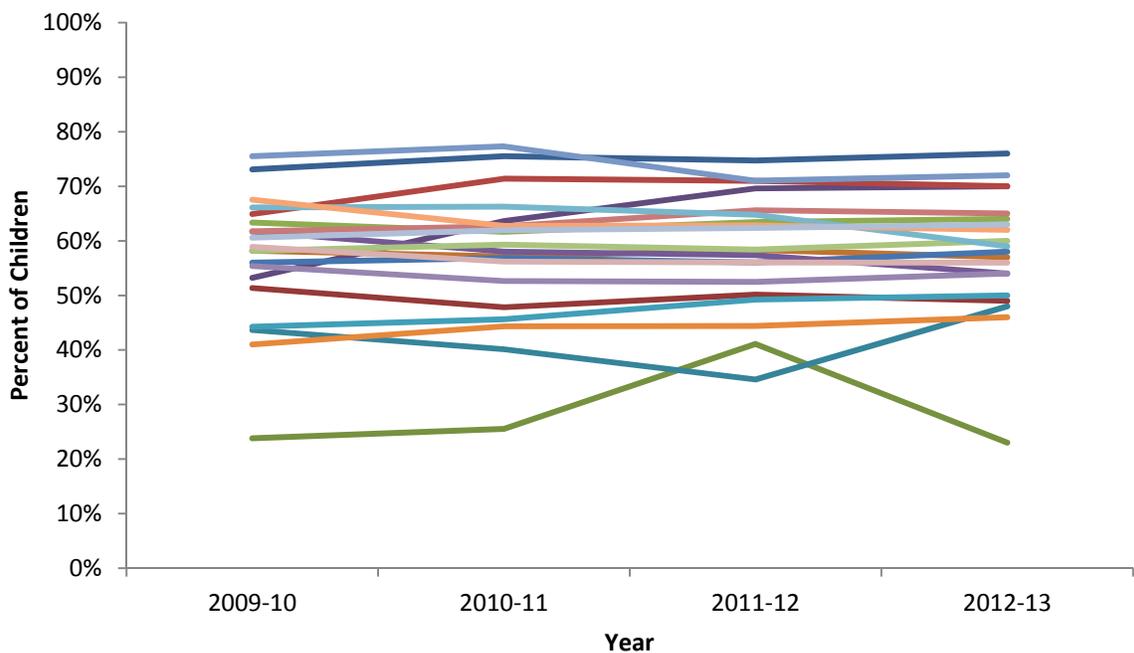
For state ECSE programs, few states had big shifts in Summary Statement 2 percentages for the knowledge and skills outcome (Exhibits 6.22–6.24). Most states showed fairly stable or slightly incremental trajectories on this outcome.

For positive social relationships and action to meet needs, about one-third of states had a moderate amount of fluctuation in Summary Statement 2 percentages. Most of the states with larger changes had fluctuations in these outcomes during 2010–11.

For state ECSE programs with Summary Statement 2, Subclaim O1 was partially supported.

Exhibit 6.22   Pattern of Exiting at Age Expectations (Summary Statement 2) Percentages for Positive Social Relationships Across Years Among State ECSE Programs Using COS Process (2009–13) (*n* = 18 states)
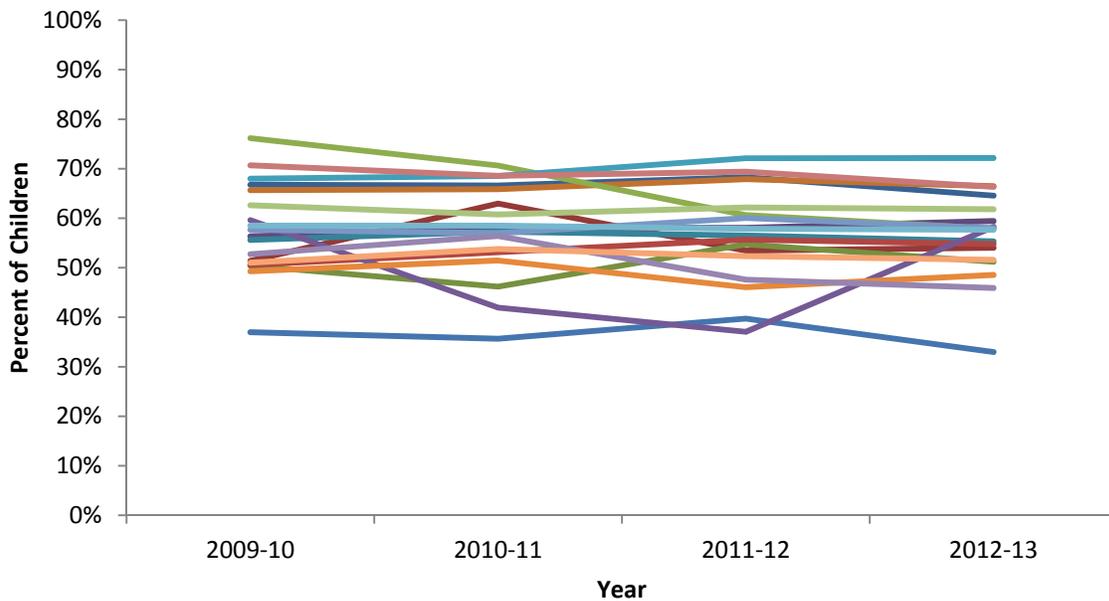
Exhibit 6.23   **Pattern of Exiting at Age Expectations (Summary Statement 2) Percentages for Knowledge and Skills Across Years Among State ECSE Programs Using COS Process (2009–13) (*n* = 18 states)**
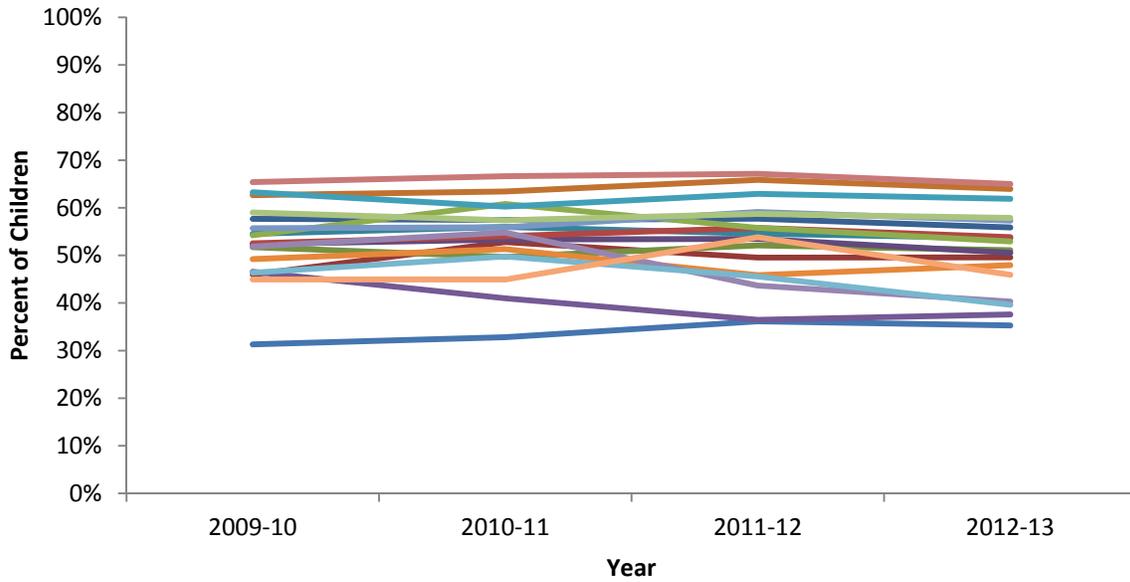
Exhibit 6.24   **Pattern of Exiting at Age Expectations (Summary Statement 2) Percentages for Action to Meet Needs Across Years Among State ECSE Programs Using COS Process (2009–13) (*n* = 18 states)**
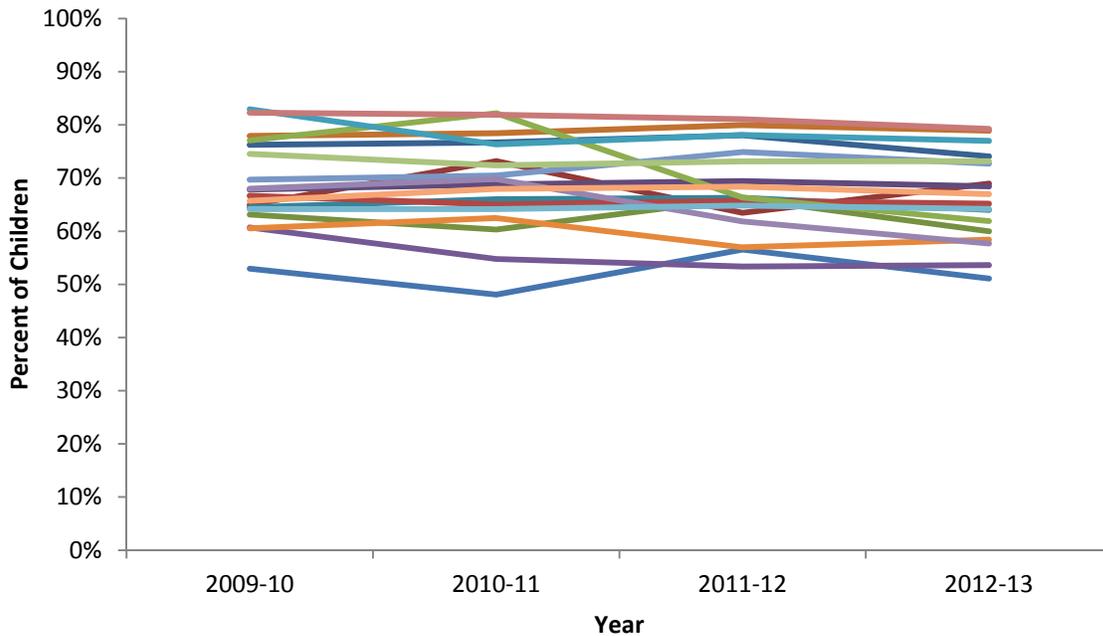
Visual inspection of stability supported Subclaim O1 in 100% of the exhibits presented here for EI across both summary statements and all three outcomes. For ECSE, 100% of the exhibits for Summary Statement 1 were supported and 33% of the exhibits for Summary Statement 2. Considered together, the visual inspection of summary statements across years in state EI and ECSE program supported Subclaim O1.

Summary, Implications, and Limitations for Study 4

Substantial evidence from the state data study suggests that the COS ratings provide information that is valid for the purpose of accountability and program improvement. Children being served in EI and ECSE were assigned the full range of COS ratings at entry and exit. Very few children entered with a rating of 1 or entered functioning at or above age expectations on all three outcomes at entry.

- Moderate correlations were found across all outcomes areas at entry and exit in both EI and ECSE.
- Children with a primary disability of autism were rated lower in positive social emotional skills than children with other disabilities, and children with a primary disability of speech or language impairment were rated higher than children with other disabilities across all outcomes.
- The year-to-year distributions of COS ratings were fairly stable especially in states with more than 1,000 exiters.
- For ECSE, all states showed correlations of .5 or higher between COS ratings at entry and exit for all outcomes.
- For EI and ECSE, all states showed variance across local programs in the percentage of children who made greater than expected growth in each of the outcomes.

One limitation of the study is that very little was known about the quality of the professional development on the COS process or the quality of the COS implementation in the participating states. Also, little was known about other activities being implemented in the states and programs like intensive quality assurance or program improvements. It is highly likely that there was variation across states in how well the COS process was being implemented. Poor training or poor implementation would weaken the relationships explored in this study of state data. Despite this likely variation, the study found support for the majority of the claims being studied.

Section 7 — Findings About Validity Claims

Earlier sections have described findings from each of the ENHANCE studies. This section presents the body of evidence from all these studies that relates to the 16 claims identified in our validity argument. We describe each validity claim, the rationale for it, and the evidence that supports or refutes it. For most claims, one or more criteria were identified to test validity. The conclusion from the review of evidence for each subclaim is summarized as *supported*, *mixed*, or *not supported*. The conclusion is mixed if more than 50% of data examined supported the subclaim, but the evidence did not meet the subclaim threshold. A summary table of the findings across all the claims and criteria is at the end of this section (Exhibit 7.13).

Claim A

Claim A: The three outcome areas reflect important outcomes for accountability and program improvement.

Rationale for Claim A

The content of the outcomes needs to be widely acknowledged as important for young children with disabilities to achieve in order for the resulting data to be useful for national and state accountability or for program improvement. It is important for stakeholders at the program, state, and national levels to consider the content of the outcomes as relevant to early intervention or early childhood special education's programmatic goals. If the content was perceived as too narrow or unrelated to program goals, the data on the outcomes could be legitimately critiqued as irrelevant to the program.

Evidence for Claim A

The evidence for the importance of the content of the three outcomes comes from the process used to derive them. The outcomes were developed through a yearlong process that involved obtaining input from multiple stakeholders, including state agency staff, local program representatives, federal policy makers, researchers, advocates, representatives from professional associations, and parents of children with disabilities. Described as "iterative social validation," the process for the identification of the three outcomes involved multiple reviews and revisions of the content with the stakeholders. The output of the stakeholder review process was released on a public website for additional comment and contributions. The general consensus from the stakeholders is that the three outcomes reflect important areas of development and learning for young children with disabilities and that they are consistent with the programmatic intent of early intervention and early childhood special education. The subsequent feedback from numerous other stakeholders over the 15 years since the outcomes were identified has been consistent with these assertions.

Conclusion for Claim A

The claim is supported.

Claim B

Claim B: The skills and behaviors described in the COS guidance materials and training resources are sufficient to enable providers to accurately describe the content of the three OSEP outcomes.

Rationale for Claim B

Using the COS process, practitioners in both EI and ECSE programs need to be able to understand the kinds of skills and behaviors associated with each of the three child outcomes in order to make ratings on children's functioning for each of the three outcomes.

Evidence for Claim B

Some support for this claim is the fact that EI and ECSE providers typically have education, training, and experience in child development that forms the basis for the content of the three OSEP child outcomes. That is, practitioners in EI and ECSE who are involved in making COS ratings work with young children and their families in a variety of ways that require that they have knowledge about typical child development across the multiple domains (e.g., conduct assessments or interview parents about children's developmental skills for determining eligibility for services, participate in the development of individualized treatment plans about promoting children's skills and behaviors). Thus, in order to work in EI or ECSE, providers have education and training in the kinds of constructs that are addressed in the three child outcomes.

Additional evidence that EI and ECSE providers understand the content of the three child outcomes comes from their responses to the online provider survey (Study 1) and from the team decision-making study showing providers working in teams to identify COS ratings (Study 3). Specifically, a majority of the providers were confident that they understand the COS process, including the content of the three OSEP child outcomes and how to make the ratings (Study 1, Exhibits 3.6 and 3.7). Providers perceived their other team members' understanding of the COS process positively but did not rate it as highly as their own understanding (Study 1, Exhibit 3.8). Teams demonstrated an ability to associate skills and behaviors with the right outcome areas both in videos (Study 3, Exhibit 5.14) and in their documentation (Study 3, Exhibit 5.42). The extent to which they consistently discussed the full breadth and depth of each outcome in their discussions varied across teams. However, the brevity of many COS team discussions, rather than a lack of understanding of the content of the outcomes, may have influenced the breadth and depth of discussion about outcomes observed in the videos.

Subclaim B1: At least 85% of the time, teams associate skills with appropriate outcomes during team discussions without major classification errors or additions. (Team Decision-Making Study)

- In the team decision-making study, nearly all teams displayed no major errors in categorizing skills with the appropriate outcome for positive social relationships (94% of teams), knowledge and skills (94% of teams), and taking action to meet needs (96% of teams) (Study 3, Exhibit 5.13.

These findings suggest that the majority of teams consistently described skills relevant to the outcome area on the COS form. Therefore, this subclaim was supported.

Subclaim B2: At least 85% of the time, teams associate skills with the appropriate outcome on the COS Form. (Team Decision-Making Study)

Among the 82% of teams that sufficiently documented skills on the Child Outcomes Summary Form, 97% of their forms had no major errors in associating children's skills with the appropriate outcome area (Study 3, Exhibit 5.42).

Therefore, this subclaim was supported.

Subclaim B3: Teams implementing the COS process will apply accurate knowledge of the full breadth of the outcomes for 85% of the outcomes rated. This will be evident by describing skills representing the core areas of each outcome in their team discussions. (Team Decision-Making Study)

- Team discussed the full breadth of the content of the outcomes in 55%, 44%, and 40% of the videos showing discussions of positive social relationships, knowledge and skills, and taking action to meet needs, respectively.
- About two-thirds of team discussions (65%) addressed the full breadth of content for at least one of the three outcome areas. Many team discussions had moderate breadth.
- A small percentage of teams showed limited breadth in discussion of positive social relationships (9%), knowledge and skills (15%), and action to meet needs (13%). Only 8% of teams had very limited breadth on two or three of the outcome areas.
- The extent to which team discussion had good breadth about the outcome and the amount of time the team spent discussing the COS in the video were positively and significantly related ($p < .05$).

A smaller than expected percentage of teams conducted discussions of limited breadth about the content of the outcomes. Those teams did not display evidence that they believed outcomes to be narrow; rather, they simply did not discuss the full breadth of outcome area content. These meetings tended to be short, which could have been the cause or the result of the less-than-full discussion. About two-thirds of the teams demonstrated a discussion with good breadth on at least one outcome area, and only 9–15% of team discussions on each outcome were of limited breadth.

The evidence for this subclaim is mixed.

Subclaim B4: Eighty-five percent of teams implementing the COS process will apply accurate knowledge to discuss the child's functioning on the outcomes with appropriate depth. This will be evident in observations of teams describing the child's skills and behaviors in sufficient depth to get a useful picture of the child's functioning on the topics described. (Team Decision-Making Study)

- Teams discussed the outcome area with appropriate depth in 46%, 37%, and 45% of the videos for positive social relationships, knowledge and skills, and action to meet needs, respectively. Many discussions had mixed depth, with appropriate depth about specific skills but not across all the different skills discussed (Study 3, Exhibit 5.16).
- Over two-thirds of the teams (69%) displayed an appropriate depth in talking about at least one of the three outcomes.
- On about one-fifth of the videos, discussions were of cursory depth or were only global (18% for positive social relationships, 22% for knowledge and skills, and 19% for action to meet needs). Only 5% of videos displayed very limited depth in team discussion of all three outcomes.

It was difficult for brief team discussions to meet the subclaim threshold for appropriate depth across all aspects of skills the team discussed. About one-fifth of the teams had limited depth to the degree that discussions were only very cursory or very global; most discussions tended to be a combination of mixed and appropriate levels of depth rather than fully appropriate depth.

The evidence for this subclaim was mixed.

Subclaim B5: At least 85% of providers in the team decision-making study indicate that "I understand the skills and behaviors included in each of the three outcomes" well or very well. (Team Decision-Making Study)

Almost all the 207 providers (97%) involved in the team decision-making study indicated that they understood the skills and behaviors included in each of the three outcomes "well or very well" (Study 3, Exhibit 5.9).

This subclaim was supported.

Subclaim B6: At least 85% of providers will indicate that it is mostly or very true that "I understand the meaning of the three outcomes." (Provider Survey)

A majority of providers participating in the online survey (85%) reported strong agreement that they understand the meaning of the three outcomes (Study 1, Exhibit 3.6).

This subclaim was supported.

Subclaim B7: At least 85% of providers will indicate that it is mostly or very true that "I know how to discuss the child's functioning in the three outcome areas with others who know the child." (Provider Survey)

A majority of providers participating in the online survey (83%) reported strong endorsement of this statement indicating that they understand the meaning of the three outcomes (Study 1, Exhibit 3.7).

The evidence for this subclaim was mixed.

Subclaim B8: At least 85% of providers will indicate it is mostly or very true that "Almost all or all professionals involved in the COS ratings understand the meaning of the three outcomes." (Provider Survey)

About two-thirds of providers (69%) reported strong endorsement of this statement about other providers' understanding the meaning of the three outcomes (Study 1, Exhibit 3.8). A stronger majority of providers (87%) endorsed the statement that it was mostly or very true that many, almost all, or all providers "understand the meaning of the three outcomes."

The evidence for this subclaim was mixed.

Exhibit 7.1   Summary of Findings for Claim B

Claim B: The skills and behaviors described in the COS guidance materials and training resources are sufficient to enable providers to accurately describe the content of the three OSEP outcomes	
Subclaim B1: At least 85% of the time, teams associate skills with the appropriate outcome without major classification errors or additions.	Supported (94%, 94%, and 96% of teams, respectively, for each outcome)
Subclaim B2: At least 85% of the time, teams associate skills with the appropriate outcome based on skills documented on the actual COS Form.	Supported (97%)
Subclaim B3: Teams implementing the COS process will apply accurate knowledge of the full breadth of the outcomes for 85% of the outcomes rated. This will be evident by describing skills representing the core areas of each outcome.	Mixed (66%)
Subclaim B4: 85% of teams implementing the COS process will apply accurate knowledge to discuss the child's functioning on the outcomes with appropriate depth. This will be evident by describing skills/functioning in sufficient depth to get a useful picture of the child's functioning on the topics described.	Mixed (69%)
Subclaim B5. At least 85% of providers in the team decision-making study indicate that "I understand the skills and behaviors included in each of the three outcomes" well or very well.	Supported (97%)
Subclaim B6: At least 85% of providers indicate that it is mostly or very true that "I understand the meaning of the three outcomes."	Supported (85%)
Subclaim B7 At least 85% of providers indicate that it is mostly or very true that "I know how to discuss the child's functioning in the three outcome areas with others who know the child."	Mixed (83%)
Subclaim B8: At least 85% of providers indicate that it is mostly or very true that "Almost all or all professionals involved in the COS ratings understand the meaning of the three outcomes."	Mixed (69%)

Conclusions for Claim B

Despite considerable variation in the level of training they had on the COS process, most providers indicated or demonstrated that they understand how children's skills and behaviors map onto the three child outcomes. The majority, but slightly fewer providers, reported being comfortable discussing the child's functioning in the three outcome areas. Although nearly all teams were able to accurately assign skills to the three outcome areas, some did not discuss the child's functioning in the outcome area with sufficient breadth or depth.

Of the eight subclaims for claim B, four subclaims were supported and four were mixed. Given that all of them were either supported or mixed, the overall claim is supported (Exhibit 7.1).

Claim C

Claim C: There is variability in children's functioning in the three outcome areas, and that variability is reflected in the COS ratings.

Rationale for Claim C

Children served in EI and ECSE enter and exit programs with a wide range of abilities and skills and also make varying degrees of progress between entry and exit. On the 7-point rating scale on the COS, the lowest rating is 1 and the highest rating is 7.

We expect that that very few children served by EI and ECSE have severe impairments in any outcome and therefore that the percentage of children with ratings of 1 should be very low. Likewise, we would expect that relatively few children requiring special education services would be functioning at a level comparable to same-age peers across all three of the functional outcomes and therefore that the percentages of children with ratings of 6 and 7 should be low at entry. Children entering with ratings of 6 or 7 on all three outcomes on the COS would usually be those who function effectively in everyday environments with assistive technology but who may not function as effectively without assistive technology supports. Another example would be children with diagnosed conditions or physical/sensory impairments who currently are functioning at age expectations but receive supportive services to keep their disability from affecting skill development and functioning.

Evidence for Claim C

Evidence for this claim comes from several studies: the extant state data study (Study 4) and the smaller samples for the child assessments study (Study 2) and the team decision-making study (Study 3

Subclaim C1: Children being served in EI and ECSE will demonstrate a full range of ratings (1 to 7) on the COS at entry and exit on each outcome. (Extant State Data, Child Assessments Study, Team Decision-Making Study)

- All the state programs whose extant data were analyzed showed the full range of COS ratings at both entry and exit.
- The full range of COS rating scores (from 1 to 7) also were recorded in both the entry sample ($n = 153$) and in the entry and exit ratings of the longitudinal sample ($n = 70$) in the child assessments study (Study 2).
- In the 73 videos of entry meetings[33] in the team decision-making study, the full range of scores was observed on each of the three outcomes (Study 3).

The full range of scores was found for the extant state data study, the child assessments study, and the team decision-making study. Therefore, this subclaim is supported.

---

[33] The distribution of the 7 COS ratings was not tested in the videos of exit meetings because only 40 exit videos were analyzed.

Subclaim C2: On each outcome, the majority of states will show that less than 10% of children being served in state EI and ECSE programs enter with ratings of 1 on the COS. (Extant State Data)

Exhibit 7.2 shows that in a majority of states, less than 10% of children entering EI or ECSE services received an entry COS rating of 1.

Exhibit 7.2   States Meeting Subclaim That Less Than 10% of Children Entering State EI or ECSE Programs Have COS Ratings of 1 at Entry (*n* = 9 states in each program)

Child Outcome	EI Programs *n* (%)	ECSE Programs *n* (%)
Positive social relationships	9 (100)	7 (78)
Knowledge and skills	6 (67)	7 (78)
Takes actions to meet needs	6 (67)	8 (89)

This subclaim is supported because a majority of the states met the subclaim for each outcome and for each program.

Subclaim C3: On each outcome, less than 10% of children served in EI and ESCE enter with ratings of 1 on the COS. These low percentages of children with COS ratings of 1 will be evident in the two research samples for the child assessments study and the team decision-making study (Child Assessments Study, Team Decision-Making Study)

In the two research samples, 5% or fewer of children participating received COS ratings of 1 at entry on each of the three outcomes.

Therefore, this subclaim is supported.

Subclaim C4: The majority of states will find that less than 15% of children served in state EI or ECSE programs are rated at or above age expectations on all three outcomes at entry. (Extant State Data)

- All eight state EI programs had less than 15% of children at age expectations on all three outcomes at entry.

- Six of the eight ECSE programs (75% of states) had less than 15% of children at age expectations on all three outcomes at entry.

This subclaim is supported because a majority of the states for both programs met the subclaim.

Subclaim C5: For each outcome, less than 15% of children will receive ratings indicating that the child is functioning at age-expected levels (i.e., ratings of 6 or 7) on all three outcomes at program entry. (Child Assessments Study, Team Decision-Making Study)

- In the entry sample of the child assessments study, 16 of 153 children (10%) had scores of 6 or 7 on all three outcomes at entry to the program.
- Data examined in the team decision-making study included videos of meetings regarding 73 children at program entry. Of these, 14% entered at age-expected levels on all three outcomes (i.e., rated as either 6 or 7 on all three outcomes at the entry meeting) (Study 3).

In both research studies, less than 15% of children entered with COS ratings of 6 or 7 on all three outcomes. Therefore, subclaim C5 is supported.

Conclusions for Claim C

Of the five subclaims for claim C, all were supported (Exhibit 7.3). Therefore, the overall claim that there is variability in children's functioning in the three outcome areas and that variability is reflected in the COS ratings is supported.

Exhibit 7.3   Summary of Findings for Claim C

Claim C: There is variability in children's functioning in the three outcome areas and that variability is reflected in the COS ratings	
Subclaim C1: Children being served in EI and ECSE will demonstrate a full range of scores (1 to 7) on the COS at entry and exit on each outcome.	Supported (100%of states))
Subclaim C2: On each outcome, the majority of states will show that less than 10% of children being served in EI and ESCE enter with ratings of 1 on the COS.	Supported (78%, 67%, 67%)
Subclaim C3: On each outcome, On each outcome, the state data will show that less than 10% of children being served in EI and ESCE enter with ratings of 1 on the COS.	Supported (4%, 5%, 3%)
Subclaim C4: The majority of states will find that less than 15% of children served in EI or ECSE will be rated at or above age expectations on all three outcomes at entry.	Supported (100%, 75%)
Subclaim C5: Less than 15% of children served in EI or ECSE will be rated at or above age expectations on all three outcomes at entry.	Supported (10%, 14%)

Claim D

Claim D: There are developmental sequences within each outcome that provide the internal structure of the COS ratings.

Rationale for Claim D

Children's skills and behaviors develop over time in predictable sequences from more simple, undifferentiated skills and behaviors to more complex and differentiated skills and

behaviors, and there are age-expected skills and behaviors that are increasingly complex with age and individual differences in functioning at any given age. The COS 7-point rating scale is based on developmental sequences, with ratings of 6 or 7 indicating that the child is showing age-appropriate functioning across settings and situations and the 1 to 5 ratings indicating varying degrees of distance from age-appropriate functioning or levels of foundational skills necessary to develop age-appropriate functioning.

Evidence for Claim D

The evidence for this claim comes from the vast literature about child development. Review articles about the history and theories of child development and a decades-long body of research support the existence of sequences of skills and behaviors in multiple domains (e.g., cognitive, social, and language development) that develop from more primitive forms to more complex forms as children age(Cairns & Cairns, 2006). For instance, developmental theories such as those of Piaget, Baldwin, and Vygotsky all posit stages or skill progressions in children's development in which earlier, less mature skills become transformed into more complex flexible skills, and this occurs in multiple domains (for example, language and motor skills) (Cairns & Cairns, 2006). More recent theoretical discussions referred to as a bioecological model define development as the "phenomenon of continuity and change in the biopsychological characteristics of human beings." And predictable developmental sequences are part of these models that examine how children's experiences and environments influence variations in rates of growth and the robustness of their skills and behaviors (e.g., ease of functional use of skills and behaviors in real-life settings and situations)(Bronfenbrenner & Morris, 2006, p. 793). Developmental theories also characterize developmental sequences as progressing from more dependent forms to more independent forms. For instance, sequences of children's social skills in interacting with adults and peers move from completely dependent and egocentric patterns or social interactions to more autonomous and reciprocal forms of interaction(Thompson, 2006).

Tools for developmental screening and assessment of young children as well as developmental milestone and growth charts are based on empirical data about sequences of skills and behaviors that are associated with age and that develop from less mature to more mature forms and functions(American Academy of Pediatrics, 2001; Division for Early Childhood, 2007, 2014).

Conclusions for Claim D

This claim is supported by the large theoretical and empirical literature about child development that attests to the existence of sequences of child development.

Claim E

Claim E: Providers of EI and ECSE services can be trained to understand and correctly apply knowledge of child development and the COS rating criteria such that a child's team will assign an accurate rating.

Rationale for Claim E

To make COS ratings that accurately reflect children's levels of functioning in each of the three outcomes, EI and ECSE providers need to have knowledge about sequences of child development, age-expected skills and behaviors during early childhood, and an understanding of the 7-point COS rating scale. If the team members involved in the COS decision-making process do not have this knowledge across them, teams will have difficulty identifying COS ratings that accurately reflect children's functioning.

COS-specific training is intended to provide an understanding of the outcomes, how to use the 7-point rating scale, and how to examine the distance of the child's functioning from age expected. However, we expect that providers of EI and ECSE services will have broader content knowledge, such as understanding sequences of child development and which skills are age expected for children of various ages because, this kind of knowledge is expected for high-quality service provision. We would expect that most providers obtain these more general skills through their educational and preservice schooling backgrounds and through opportunities for continuous learning throughout their careers. This claim examines whether or not there is evidence that providers possess some key skills needed to identify COS ratings effectively.

Evidence for Claim E

Evidence about EI and ECSE providers' knowledge and requisite skills for COS decision-making comes from multiple sources. The online provider survey (Study 1) yielded information about whether EI and ECSE providers can understand and correctly apply knowledge of child development and the COS rating criteria to make accurate ratings. The team decision-making study (Study 3) examined whether providers demonstrate these skills during videos of COS meetings. We would expect to see that a majority of the providers are confident that they understand the COS process, including the content of the three OSEP child outcomes and how to make the ratings (Study 1, Exhibits 3.6 and 3.7). We also would expect that providers' perceptions of their other team members' understanding of the COS process would be generally positive (Study 1, Exhibit 3.8).

The team decision-making study data found considerable variability in COS process implementation across teams using the COS rating process. It may be that teams that work together regularly discuss child outcomes in a shorthand way because the team members share a significant amount of knowledge. Nevertheless, for those videos with sufficient information, we would expect that the data would indicate that the COS process is being implemented sufficiently well to yield accurate data.

Subclaim E1: In 85% of the videos in the team decision-making study, teams implementing the COS process will not have major instances of inappropriate age anchoring of skills described. (Team Decision-Making Study)

- In 91% of videos ($n = 87$) where teams' age anchoring was observed, coders did not observe any major errors in age anchoring (Study 3, Exhibit 5.23).
- In 84% of videos ($n = 87$) where teams' age anchoring was observed, coders did not observe either major or minor errors in age anchoring.
- Only 77% of the videos showed teams age anchoring skills in such a way that the accuracy of the age anchoring could be coded. For the other videos, it was impossible to identify whether teams made any age-anchoring errors.

Based on the strong percentage of videos where age anchoring was observed without any major errors, this subclaim was supported.

Subclaim E2: At least 85% of providers will indicate mostly/very true that "I understand what is age-expected functioning in each of the three outcome areas." (Provider Survey)

A majority of providers participating in the online survey (89%) reported strong agreement that they understand the meaning of the three outcomes (Study 1, Exhibit 3.6).

This subclaim was supported.

Subclaim E3: At least 85% of providers will indicate that it is mostly/very true that "I know how to compare the child's functioning to age-expected functioning." (Provider Survey)

A majority of providers participating in the online survey (91%) reported strong agreement that they understand how to compare a child's functioning to age-expected functioning (Study 1, Exhibit 3.7).

This subclaim was supported.

Subclaim E4: At least 85% of providers will indicate that almost all/all providers "understand what is age-expected functioning in each of the three outcome areas." (Provider Survey)

- A majority of providers participating in the online survey (74%) reported that almost all or all other providers understand what is age-expected functioning in each of the three outcome areas (Study 1, Exhibit 3.8).
- A stronger majority of providers (91%) agreed that many/almost all/all providers "understand what is age-expected functioning in each of the three outcome areas."

The evidence for this subclaim was mixed.

Subclaim E5: At least 85% of providers will indicate that they understand well or very well "the definition of the 7 rating points." (Team Decision-Making Study)

A majority of providers participating in the team decision-making study (94%) reported that they understood the definition of the 7 rating points either well or very well (Study 1, Exhibit 3.6).

This subclaim was supported.

Subclaim E6: At least 85% of providers will indicate that it is mostly or very true that "I know how to identify how the child uses his/her skills to perform meaningful, everyday tasks." (Provider Survey)

A majority of providers participating in the online survey (90%) reported strong endorsement of this statement (Study 1, Exhibit 3.7).

This subclaim was supported.

Subclaim E7: At least 85% of providers will indicate that it is mostly or very true that "Almost all or all professionals involved in the COS ratings know the difference between functional behaviors and discrete skills." (Provider Survey)

- Over half the providers participating in the online survey (59%) reported that all or almost all of their other team members "know the difference between functional behaviors and discrete skills" (Study 1, Exhibit 3.8).
- A larger majority of providers (82%) endorsed that many, almost all, or all providers "know the difference between functional behaviors and discrete skills."

The evidence for this subclaim was mixed.

Subclaim E8: At least 85% of providers will indicate that "I understand the degree to which different skills and behaviors are age appropriate" well or very well. (Team Decision-Making Study)

A majority of providers participating in the team decision-making study (97%) indicated that "I understand the degree to which different skills and behaviors are age appropriate" (Study 1, Exhibit 3.6).

This subclaim was supported.

Subclaim E9: For at least 85% of outcomes rated, teams implementing the COS process will not have a misapplication/misunderstanding of rating criteria based on problems with sequences or with anchoring age-level skills. (Team Decision-Making Study)

The majority of teams did not have misunderstandings/ misapplications of the ratings based on problems with age anchoring. The number of teams with misunderstandings varied slightly across the three outcomes, with 90%, 98%, and 100% of teams avoiding misapplications when discussing positive social relationships, knowledge and skills, and action to meet needs, respectively.

For each of the three outcomes, the percentage of teams that did not have misunderstandings exceeded the threshold for the subclaim. Therefore, this subclaim was supported.

Subclaim E10: For at least 85% of outcomes rated, teams implementing the COS process will not have a misapplication/misunderstanding of rating criteria based on issues with the criteria (not age anchoring problems). (Team Decision-Making Study)

The majority of teams did not have misunderstandings in applying rating criteria to decisions. The amount of misunderstandings varied across the three outcomes, with 77%, 88%, and 86% of teams avoiding misunderstanding in discussions about positive social relationships, knowledge and skills, and action to meet needs, respectively.

The subclaim was met for two of the three outcome areas. Therefore, the evidence for this subclaim was mixed.

Subclaim E11: Teams implementing the COS process will assign ratings consistent with those assigned by an external coder trained in the COS process for 85% of the outcomes rated. (Team Decision-Making Study)

For 89% of positive social relationships discussions, 94% of knowledge and skills discussions, and 94% of taking action to meet needs discussions on the videos, the teams' ratings were within one point of the coder's ratings (Study 3, Exhibit 5.X30 and 5.31). These data are based on videos where there was enough information provided about the child's functioning for the coder to identify an appropriate rating.

This subclaim was supported.

Subclaim E12: At least 85% of providers will indicate that it is mostly or very true that "I understand how to apply the criteria for each of the 7 rating points." (Provider Survey)

A majority of providers participating in the online survey (79%) reported strong endorsement of this statement (Study 1, Exhibit 3.6).

The evidence for this subclaim was mixed.

Subclaim E13: At least 85% of providers will indicate it is mostly or very true that almost all or all providers "understand how to apply the criteria for each of the 7 rating points." (Provider Survey)

- About two-thirds of providers participating in the online survey (62%) reported strong endorsement of this statement (Study 1, Exhibit 3.6).
- A stronger majority (86%) of providers indicated that many, almost all, or all providers "understand how to apply the criteria for each of the 7 rating points."

The evidence for this subclaim was mixed.

Exhibit 7.4   Summary of Findings for Claim E

Claim E: Providers of EI and ECSE services can be trained to understand and correctly apply knowledge of child development and the COS rating criteria such that a child's team will assign an accurate rating	
Subclaim E1: In 85% of the videos, teams implementing the COS process will not have major instances of inappropriate age anchoring of skills described.	Supported (91%)
Subclaim E2: At least 85% of providers indicate that it is mostly or very true that "I understand what is age-expected functioning in each of the three outcome areas."	Supported (89%)
Subclaim E3: At least 85% of providers indicate that it is mostly or very true that "I know how to compare the child's functioning to age-expected functioning."	Supported (91%)
Subclaim E4: At least 85% of providers indicate that almost all or all providers "understand what is age-expected functioning in each of the 3 outcome areas."	Mixed (74%)
Subclaim E5: At least 85% of providers will indicate that they understand well or very well "the definition of the 7 rating points."	Supported (94%)
Subclaim E6: At least 85% of providers indicate that it is mostly or very true that "I know how to identify how the child uses his/her skills to perform meaningful, everyday tasks."	Supported (90%)
Subclaim E7: At least 85% of providers indicate that almost all or all professionals involved in the COS ratings "know the difference between functional behaviors and discrete skills."	Mixed (59%)
Subclaim E8: At least 85% of providers will indicate that "I understand the degree to which different skills and behaviors are age appropriate" well or very well.	Supported (97%)
Subclaim E9: For at least 85% of outcomes rated, teams implementing the COS process will not have a misapplication/misunderstanding of rating criteria based on problems with sequences or with anchoring age-level skills	Supported (90%, 98%, 100%)
Subclaim E10: For at least 85% of outcomes rated, teams implementing the COS process will not have a misapplication/ misunderstanding of rating criteria based on issues with the criteria (not age anchoring problems)	Mixed (77%, 88%, 86%)
Subclaim E11: Teams implementing the COS process will assign ratings consistent with those assigned by an external coder trained in the COS process for 85% of the outcomes rated.	Supported (89%, 94%, 94%)
Subclaim E12: At least 85% of providers indicate that it is mostly or very true that "I understand how to apply the criteria for each of the 7 rating points."	Mixed (79%)
Subclaim E14: At least 85% of providers indicate that it is mostly or very true that almost all or all providers "understand how to apply the criteria for each of the 7 rating points."	Mixed (62%)

Conclusions for Claim E

Criteria identified in Claim E examined the extent to which providers can learn the skills needed for the COS process. These criteria were tested in samples with highly variable levels of provider training about the COS process. The training the providers received was not uniform across states, and it was not provided or controlled by the research team.

Of the 14 subclaims for Claim E, all were either supported or the evidence was mixed (9 supported and 5 mixed) (Exhibit 7.4). Most of the criteria that were mixed involved providers' appraisals of their co-workers' level of understanding or skills needed for the COS process. Criteria where providers rated their own understanding or competence or where a coder examined features of COS meetings on video were generally supported. Taken together, this evidence supports Claim E.

Claim F

Claim F: Functioning in one outcome area is related to functioning in another outcome area.

Rationale for Claim F

Although children's development is often divided into developmental domains for descriptive, assessment, or scientific purposes, children's functioning requires integrating skills across various domains to achieve desired tasks. The development of skills in one domain often influences how children undertake things and influences the development of skills in another domain(National Association for the Education of Young Children, 2009; National Research Council and Institute of Medicine, 2000b). This integration is especially pronounced in early childhood as foundational skills in language, movement, and conceptual understanding open up new possibilities for how children act on their worlds and what additional skills or behaviors they initiate, practice, and incorporate into their functioning. Evidence supports the extensive neural interconnections that emerge during this period and facilitate integration(Siegel, 2012).

The Child Outcomes Summary process produces ratings of children in three broad outcome areas that cut across developmental domains and reflect a child's functioning to achieve specific aims. As a result of the nature of these outcomes and the interrelatedness of child development, we expect that a team's rating of a child's functioning in one outcome area is likely to be at least moderately related to the child's functioning in other outcome areas. ENHANCE examined whether the COS data reflect the interrelatedness in functioning expected during early childhood development.

Evidence for Claim F

To examine the extent that COS ratings reflect interrelatedness in functioning across outcome areas, we examined correlations between COS ratings for the three child outcomes using the statewide samples from the extant state data study, the smaller samples from the child

assessments study, and the team decision-making study. We expected that the COS ratings would show moderately strong correlations (stronger than .5) across outcomes at entry and at exit for both the state samples and the smaller program samples.

Subclaim F1: COS ratings will show at least moderately strong correlations (.5 or greater) across outcomes at entry and at exit. (Extant State Data Study, Child Assessments Study, Team Decision-Making Study)

- The extant data study found that all nine state EI and nine state ECSE programs met this subclaim. Correlations between the COS ratings on the three outcomes were stronger than .5 at both entry and exit in all the states that provided data (Study 4). Across the three outcome pairs, correlation ranges at entry were.70–.83 for ECSE and .55–.80 for EI; at exit correlation ranges were .70–.87 for ECSE and .74–.88 for EI.
- Correlations of COS ratings between the three outcomes from the child assessments study ranged from .67 to .70 at entry (*n* = 153) and .80 to .83 at exit (*n* = 70) (Study 2, Exhibits 4.22 and 4.24).
- Correlations of COS ratings between the three outcomes from the team decision-making study ranging from .75 to .81 at entry (*n* = 73) and .75 to .80 at exit (*n* = 40) (Study 3, Exhibits 5.5 and 5.6).

This subclaim was supported.

Conclusions for Claim F

All three studies found correlations between COS ratings across the three outcomes to be stronger than .5. These findings provide support for the subclaim and claim. These findings suggest that there are relationships between COS ratings on the three outcome areas that are consistent with research demonstrating that developmental domains are interrelated and that functioning in one domain influences functioning on other domains. Claim F is supported.

Claim G

Claim G: COS ratings in the corresponding outcomes are moderately related with the social emotional (outcome 1), cognitive (outcome 2), communication (outcome 2), and adaptive (outcome 3) domain scores of assessment tools.

Rationale for Claim G

COS ratings are intended to reflect the child's level of functioning on the three global outcome areas. The outcomes are broad and focus on the child's ability to integrate skills across a variety of domains to accomplish meaningful everyday tasks. The Early Childhood Outcomes Center identified the different kinds of skills and behaviors most closely related to each of these outcomes and crosswalked items and/or domains from a number of commonly used assessment tools to the outcomes (see http://ectacenter.org/eco/pages/crosswalks.asp). This process specified which assessment tool domains are most commonly associated with each outcome. Those

relationships are noted above in the claim. ECO Center guidance also provided additional information about which outcomes are related to each domain and considered how communication skills cut across all three outcomes. Although the content of tools can be cross-walked to the outcomes, the nature of the outcomes is such that the items on existing standardized assessment tools and their derived domain scores do not map perfectly or one to one onto the three functional outcomes. For example, although the motor domain is most closely linked to outcome 3 because children use motor skills to get their needs met, the lack of motor skills does not necessarily mean a child cannot get his or her needs met. Children who use walkers or wheelchairs also can get from place to place to get what they want although these children would score low in a motor domain on a typical assessments. Some states that are using only traditional assessments for federal outcomes reporting use both adaptive and motor domains for computing progress on outcome 3. However, as explained with the example above, interpretation of data derived from motor scores is problematic because the outcome addresses mobility and incorporates use of assistive devices and a motor domain score does not. Generally, we present data on the relationship between outcome 3, action to meet needs, and the adaptive domains, but we sometime also may report data on the motor domain.

Evidence for Claim G

Subclaim G1: The majority of COS ratings will show moderately strong correlations (greater than or equal to .5) with social emotional (outcome 1), cognitive (outcome 2), communication (outcome 2), and adaptive (outcome 3) domain scores of assessment tools. (Child Assessments Study)

Exhibit 7.5 shows the correlations between COS ratings and the two different assessment tools at entry in the child assessments study. Correlations in the shaded boxes are those expected to be related to each other.

Exhibit 7.5   Correlations Between COS Ratings and Domain Scores on Assessment Tools in Child Assessments Study Entry Sample (*n* = 153)

	Entry Sample	Positive Social Relationships	Knowledge and Skills	Action to Meet Needs
COS	Positive social relationships			
	Knowledge and skills	.70		
	Action to meet needs	.67	.67	
BDI-2	Personal social	.46	.43	.38
	Communication	.34	.46	.30
	Cognitive	.43	.41	.35
	Motor	.35	.37	.46
	Adaptive	.46	.41	.46
Vine-land-II	Socialization	.28	.29	.31
	Communication	.32	.34	.32
	Daily living skills	.43	.32	.40
	Motor skills	.33	.30	.50

The findings show that

- One of the two correlations of skills with positive social relationships (50%) is in the expected moderate range, with rounding.
- One out of three of the correlations for knowledge and skills (33%) was in the moderate range.
- One-half of the correlations between adaptive/daily living skills domains and action to meet needs (50%) were in the expected moderate range.
- Both the correlations between motor domains and action to meet needs (100%) were in the expected moderate range.
- The strength of the correlations differed for the COS with the different assessment tools.
  - Four out of five of the correlations with the BDI-2 (100%) were in the expected moderate range after rounding.
  - Only one out of four of the correlations with the Vineland-II (25%) were in the expected moderate range.
- Correlations that were not in the moderate range were consistently in the predicted direction, often quite near but not quite meeting the cutoff threshold.

The mixed nature of these findings both with diverse assessment tools and across outcomes, as well as the proximity of correlation values to the cutoff threshold in instances where the subclaim was not met, indicates that the evidence for subclaim G1 is mixed.

Subclaim G2: Mean assessment tool scores on social emotional (outcome 1), cognitive (outcome 2), communication (outcome 2), and adaptive (outcome) 3 will increase at each increased level of COS ratings, forming a stair-step pattern. (Child Assessments Study)

Mean assessment tool scores of children with different entry COS ratings from the child assessment study are summarized by outcome (Study 2, Exhibits 4.19–4.21). The sample size of 153 was too small to create reliable mean estimates at all 7 points on the scale, so the data described below show mean assessment tool scores at three levels of the COS, children with ratings of 1–3, 4–5, and 6–7.

Overall, across the three assessment tools examined in relation to COS ratings on the outcomes, mean assessment tool scores for children with different COS ratings showed the expected stair-step pattern of increasing mean scores in 11 out of 12 of the mean scores calculated (92%) when COS ratings were clustered into three groups (1–3, 4–5, 6–7).

Positive Social Relationships

- Across the assessment tools, for positive social relationships the stair-step pattern held for the three levels of COS ratings, with higher mean scores for children with higher COS ratings (Study 2, Exhibit 4.19).

- Similarly, across the three assessment tools, for positive social relationships the majority of the 7 stair- step patterns held at each level (Appendix K, Exhibit K-1).

Knowledge and Skills

- Across the assessment tools, for knowledge and skills the stair-step pattern held for the three levels of COS ratings, with higher mean scores for children with higher COS ratings (Study 2, Exhibit 4.20).

- Similarly, across the three different assessment tools, for knowledge and skills the majority of the 7 stair-step patterns held at each level (Appendix K, Exhibit K-2).

Action to Meet Needs

- Across the three different assessment tools, for action to meet needs the stair-step pattern held for the three levels of COS ratings, with higher mean scores for children with higher COS ratings (Study 2, Exhibit 4.21).

- Similarly, across the three different assessment tools, for action to meet needs the majority of the 7 stair-step patterns holds at each level (Appendix K, Exhibit K-3).

Exhibit 7.6   Summary of Findings for Claim G

Claim G: COS ratings in the corresponding outcomes are moderately related with the social emotional (outcome 1), cognitive (outcome 2), communication (outcome 2), and adaptive (outcome 3) domain scores of assessment tools	
Subclaim G1: COS ratings will show moderately strong correlations (greater than or equal to .5 with rounding) with social emotional (outcome 1), cognitive (outcome 2), communication (outcome 2), and adaptive (outcome 3) domain scores of assessment tools.	Mixed 　BDI-2 (80% of the correlations) 　Vineland-II (25%) 　Positive social relationships (50%) 　Knowledge and skills (33%) 　Action to meet needs (75%)
Subclaim G2: Mean assessment tool scores on social emotional (outcome 1), cognitive (outcome 2), communication (outcome 2), and adaptive (outcome 3 will increase at each increased level of COS ratings.	Supported (92%)

Conclusions for Claim G

Evidence suggests moderate relationships exist between the COS ratings and domain scores on other assessment tools. Correlations were slightly below the predicted value of .5. Evidence also supported the predicted stair-step patterns for mean assessment tool scores and associated COS ratings from low to high COS ratings (i.e., lower mean assessment tool scores for low COS ratings than for higher COS ratings) That is, across the three assessment tools examined in relation to COS ratings on the outcomes, mean assessment tool scores for children with different COS ratings showed the expected patterns of increasing mean scores, a stair-step pattern, in 11 out of 12 of the mean scores calculated (92%) when COS ratings were clustered into three groups (1–3, 4–5, 6–7). The only exception was with the Vineland-II communication where mean scores on communication were similar for COS rating groups 4–5 and 6–7.

Taken together, the evidence for Claim G is mixed (Exhibit 7.6).

Claim H

Claim H: COS ratings will be related to the type and severity of the child's disability.

Rationale for Claim H

Children receiving EI and ECSE services have a wide range of different types of disabilities and a wide range of functioning levels. Children with some types of disabilities show more significant levels of functioning (e.g., children with autism are often more impaired and children with primarily speech language impairments are often higher functioning). In addition, for some types of disabilities, functioning is often more impaired for some outcomes than others (e.g., children with a diagnosis of autism may have greater deficits in functioning of positive social relationships than the other two child outcomes).

Evidence for Claim H

Data from the child assessment study and the extant state data study were used to examine this claim. Data on functional levels for children with different types of disabilities were compared with the expectation that these differences would be reflected in COS ratings.

Using entry data from the child assessment study, we compared the ABILITIES Index, a global measure of functioning, with COS ratings for children with higher versus lower ABILITIES Index scores (n = 153). The same analysis was repeated with entry data from the team decision-making study (n = 73).

Extant data from the two states that provided COS ratings by primary disability for ECSE also were used to test this claim. Specifically, COS ratings for ECSE children in two primary disability categories, autism and speech or language impairment, were compared with COS data for children with all other disabilities. Data from EI could not be used to examine claim H because EI programs do not use the federal reporting categories for disabilities.

Subclaim H1: Children in the lowest third of ratings on the ABILITIES Index at entry will have significantly lower COS ratings than children in the higher two-thirds of the ABILITIES Index. (Child Assessments study, Team Decision-Making Study)

As shown in Exhibit 7.7, for each of the three child outcomes the mean COS ratings were higher for those children with higher than lower total scores on the ABILITIES Index.

Exhibit 7.7   Mean COS Ratings for Children with Higher Than Lower ABILITIES Index Scores in Child Assessments Study and Team Decision-Making Study

Mean Ratings (SD)	Positive Social Relationships		Knowledge and Skills		Action to Meet Needs	
	ABILITIES Index* < 100 (*n* = 50)	ABILITIES Index 100 or Higher (*n* = 103)	ABILITIES Index* < 100 (*n* = 50)	ABILITIES Index 100 or Higher (*n* = 103)	ABILITIES Index* < 100 (*n* = 50)	ABILITIES Index 100 or Higher (*n* = 103)
Child Assessments Study						
Entry sample (*n* = 153)	3.4 (1.51)	5.1* (1.35)	3.2 (1.57)	4.6* (1.28)	3.2 (1.42)	4.9* (1.22)
Team Decision-Making Study						
Entry meetings (Total *n* = 73; ABILITIES < 100 *n* = 23, 100 or higher *n* = 50)	3.26 (1.54)	5.46*** (1.22)	3.13 (1.74)	4.94*** (1.19)	3.26 (1.25)	5.04*** (1.11)
Exit meetings (Total *n* = 40; ABILITIES < 100 *n* = 11, 100 or higher *n* = 29)	4.55 (1.13)	5.83* 1.26)	4.27 (0.90)	5.55* (1.18)	4.18 (1.17)	6.03*** (1.12)

* p < .01, ** p < .001, *** p < .0001.

Note: ABILITIES Index scores were reversed such that lower scores represent lower functioning and higher scores represent higher functioning.

These data show that the COS ratings differentiate children with higher and lower levels of overall developmental functioning. Therefore, subclaim H1 was supported.

Subclaim H2: Children receiving ECSE services with a primary disability type of "other: autism/spectrum disorder" will have significantly lower COS ratings on the positive social relationships outcome than children with other diagnoses. Also, children with autism/spectrum disorder will have lower COS ratings on positive social relationships than on knowledge and skills and action to meet needs compared with children who have other disabilities. (Extant Data Study)

Subclaim H2 was tested using extant data from two states. The ratings of children with a primary diagnosis of autism were compared with ratings for children with all other primary disabilities. To meet subclaim H2, two data patterns were expected. First, children with a primary disability of autism should be rated lower than children with other disabilities in positive social relationships. The analyses showed that

- In state 1, 99% of children with a primary disability of autism were rated below age expectations in positive social relationships compared with 75% of children with other primary disabilities.
- In state 2, 99% of children with a primary disability of autism were rated below age expectations in positive social relationships compared with 90% of children with other primary disabilities.

These results are in the expected directions in both states and lend support to subclaim H2.

Second, children with a primary disability of autism were expected to be rated lower in positive social relationships than in knowledge and skills or in actions to meet needs.

- In state 1, 99% of children with a primary disability of autism were rated below age expectations in positive social relationships compared with 92% rated below age expectations in knowledge and skills and 92% rated below age expectations in actions to meet needs.
- In state 2, 99% of children with a primary disability of autism were rated below age expectations in positive social relationships compared with 99% rated below age expectations in knowledge and skills and 96% rated below age expectations in actions to meet needs.

Findings show results in the expected directions but not with strong differences. Results are stronger in state 1 than in state 2. Taken together, the evidence for subclaim H2 is mixed.

Subclaim H3: Children receiving ECSE services with a primary disability of speech-language impairment will have significantly higher COS ratings on each outcome than children with other primary disabilities. (Extant Data Study)

To meet subclaim H3, children who had a primary disability of speech or language impairment should have been rated higher at entry for each outcome relative to a comparison group of children with all other primary disabilities. We expected that fewer children with a

speech-language disability would have scores below age-level expectations on each outcome. This subclaim was examined in the same two states used for subclaim H2.

Data for both states showed that fewer children were rated below age expectations for each outcome in the speech or language impairment group compared with children who had other primary disabilities for all three outcomes ($p < .001$)(Study 4, Exhibit 6.4). This trend was more pronounced for state 1 than state 2, however difference in both cases were statistically significant. Subclaim H3 was supported.

Subclaim H4: The reversed ABILITIES Index total score (i.e., higher number represents higher functioning level) should be moderately correlated (.5 or greater) with COS ratings at entry. (Child Assessments Study, Team Decision-Making Study)

- In the child assessments study, correlations between COS ratings and the reversed ABILITIES Index scores were .54 for positive social relationships, .53 for knowledge and skills, and .59 for taking action to meet needs in the entry sample ($n = 153$).
- In the team decision-making study, correlations between COS ratings and the reversed ABILITIES Index scores were .59 for positive social relationships, .65 for knowledge and skills, and .71 for taking action to meet needs among entry videos ($n = 73$).

These data indicate moderate to strong relationships between disability ratings of the child's overall functioning and the child's COS rating. Evidence from two different studies suggests that subclaim H4 was supported.

Exhibit 7.8   Summary of Findings for Claim H

Claim H: COS ratings will be related to the type and severity of the child's disability	
Subclaim H1: Children with lowest one-third of the ABILITIES Index ratings at entry will have significantly lower COS ratings than children in the higher two-thirds of the ABILITIES Index.	Supported
Subclaim H2: ECSE children with a disability type of" other: autism/ spectrum disorder" will have significantly lower COS ratings on outcome 1 than children with other diagnoses; children with autism/spectrum disorder also will have lower scores on outcome 1 than on outcomes 2 and 3.	Mixed
Subclaim H3: ECSE children with speech-language impairment as a primary disability type will have significantly higher COS ratings on each outcome than children with other diagnoses.	Supported
Subclaim H4: Reversed ABILITIES Index total scores should be moderately correlated (.5 or greater) with COS ratings at entry.	Supported

Conclusions for Claim H

Data across three different studies show relationships between a child's level of functioning and COS ratings, with expected patterns of associations for children with different kinds of primary disability diagnoses. Three of the four criteria for claim H were supported, and the remaining subclaim was mixed (Exhibit 7.8); therefore, claim H was supported.

Claim I

Claim I: COS ratings will not be related to the composition of the team, the particular assessment tool used, or child characteristics such as gender or race/ethnicity among children with the same types and severities of disabilities.

Rationale for Claim I

Teams are expected to identify the appropriate rating to describe a child's functioning by synthesizing all available information from multiple sources about a child's development and his or her current functioning across settings and situations. Therefore, differences in ratings are not expected between children who have similar levels of functioning but differ in other nonrelevant characteristics, such as gender or race/ethnicity. Likewise, if teams are implementing the COS with fidelity to the rating criteria, ratings should not be influenced by which team members are present or the number of people on the team.

Evidence for Claim I

COS ratings were expected to be related to the child's functioning. Therefore, tests of Claim I examined relationships between each predictor variable and COS ratings after taking into account the child's level of functioning. In both the child assessments study and the team decision-making study, information about the child's disability or related conditions was limited. The best information available to approximate the child's level of functioning was the ABILITIES Index. The same analysis was repeated using entry COS ratings from both the child assessments study and the team decision-making study. The subgroup of items on the ABILITIES Index most directly related to each outcome area were used as the covariate in regressions.

Subclaim I1: Parent participation on the team will not significantly predict COS ratings after controlling for relevant scores on subareas of the ABILITIES Index. (Child Assessments Study, Team Decision-Making Study)

- For all three outcome areas, no statistically significant relationship was found between the presence of a parent or family member in the COS team discussion and entry COS ratings in the child assessments study. (Study 2, Exhibit 4.18).
- For all three outcome areas, no statistically significant relationship was found between the presence of a parent or family member in the COS team discussion and entry COS ratings in the team decision-making study. (Study 3, Exhibit 5.33).

On the basis of the evidence in the child assessments study and the team decision-making study, this subclaim was supported.

Subclaim I2: The number of members involved in the COS team decision will not significantly predict COS ratings after controlling for relevant scores on subareas of the ABILITIES Index. (Child Assessments Study, Team Decision-Making Study)

- For all three outcome areas, no statistically significant relationship was found between the number of members participating in the COS team discussion and entry COS ratings in the child assessments study. (Study 2, Exhibit 4.18).
- For all three outcome areas, no statistically significant relationship was found between the number of members participating in the COS team discussion and entry COS ratings in the team decision-making study. (Study 3, Exhibit 5.40).

On the basis of the evidence in the child assessments study and the team decision-making study, this subclaim was supported.

Subclaim I3: The child's gender will not significantly predict COS ratings after controlling for relevant scores on subareas of the ABILITIES Index. (Child Assessments Study, Team Decision-Making Study)

- For all three outcome areas, no statistically significant relationship was found between the child's gender and entry COS ratings in the child assessments study. (Study 2, Exhibit 4.18).
- For all three outcome areas, no statistically significant relationship was found between the child's gender and entry COS ratings in the team decision-making study. (Study 3, Exhibit 5.40).

On the basis of the evidence in the child assessments study and the team decision-making study, this subclaim was supported.

Subclaim I4: The child's race/ethnicity will not significantly predict COS ratings after controlling for relevant scores on subareas of the ABILITIES Index. (Child Assessments Study, Team Decision-Making Study)

- For positive social relationships and knowledge and skills, in the child assessments study no statistically significant relationship was found between race/ethnicity and entry COS ratings. (Study 2, Exhibit 4.18).
- For taking appropriate action to meet needs, in the child assessments study differences were found across racial ethnic groups. (Study 2, Exhibit 4.18):
  - No statistically significant differences were observed in the COS ratings for children with Hispanic or "other" race/ethnicity categories compared with Caucasian children.
  - Being an African American child was related to lower COS ratings on action to meet needs ($p < .05$) after controlling for subarea scores on the ABILITIES Index.
- For all three outcomes, in the team decision-making study no statistically significant relationship was found between any of the child's race/ethnicity categories and entry COS ratings. (Study 3, Exhibit 5.40).

Evidence in the child assessments study supports this subclaim for two of the three outcomes. On the third outcome, the subclaim was supported for two of the three the racial/ethnic comparisons considered. Therefore, taken together, this subclaim was supported.

Subclaim I5: The specific assessment tool referenced by teams in the COS process will not significantly predict COS ratings after controlling for relevant scores on subareas of the ABILITIES Index and EI/ECSE program participation. (Team Decision-Making Study)

To test this subclaim using data from the team decision-making study, we examined a regression predicting COS ratings at entry on each outcome from the four most common assessment tools used during the COS process and a collective referent of any other assessment tool used. The specific assessment tools[34] included were the Battelle Developmental Inventory ($n = 14$), Creative Curriculum ($n = 16$), the Preschool Language Scale (PLS) ($n = 10$), and the Infant-Toddler Developmental Assessment (IDA) ($n = 12$). Because in some cases the assessment tool was systematically related to EI versus ECSE program status (e.g., the IDA is used only in EI, the PLS is used only in ECSE), EI/ECSE program status as well as scores on relevant subareas of the ABILITIES Index were included as covariates.

Across all three outcomes, none of the four tested assessment tools showed a statistically significant relationship to COS ratings at entry.

These results showed that COS ratings are not influenced by specific assessment tools that teams used for information about the child's functioning. Since all four assessment tools showed the expected relationships across all three outcomes, this subclaim was supported.

---

[34] All versions of the same assessment tool were collapsed for this analysis. For instance, teams referencing the BDI and the BDI-2 were all counted in the BDI group.

Exhibit 7.9   Summary of Findings for Claim I

Claim I: COS ratings will not be related to the composition of the team, the particular assessment tool used, or child characteristics such as gender or race/ethnicity among children with the same types and severities of disabilities	
Subclaim I1. Parent participation on the team will not significantly predict COS ratings after controlling for relevant scores on subareas of the ABILITIES Index.	Supported
Subclaim I2. The number of members involved in the COS team decision will not significantly predict COS ratings after controlling for relevant scores on subareas of the ABILITIES Index.	Supported
Subclaim I3. The child's gender will not significantly predict COS ratings after controlling for relevant scores on subareas of the ABILITIES Index.	Supported
Subclaim I4. The child's race/ethnicity will not significantly predict COS ratings after controlling for relevant scores on subareas of the ABILITIES Index.	Mixed
Subclaim I5. The specific assessment tool referenced by teams in the COS process will not significantly predict COS ratings after controlling for relevant scores on subareas of the ABILITIES Index and EI/ECSE program participation.	Supported

Conclusions for Claim I

Support was found for four of the five subclaims for claim I (Exhibit 7.9). In data from two different studies, COS ratings at entry were not related to parent participation, number of team members, the child's gender, the child's race/ethnicity, and the specific assessment tool used during the COS process after taking into account the child's level of functioning. Taken together, the data support claim I.

Claim J

Claim J: COS entry scores will be similar for programs and regions serving similar populations.

Rationale for Claim J

If the COS data are valid, then similar types of children (e.g., with similar disabilities and severity levels of functioning) should score similarly. At an aggregate level, this would mean that programs or regions serving similar populations also would have similar aggregate scores at entry to the program. We would not expect similarities at exit because the programs might be differentially effective

Evidence for Claim J

We were not able to obtain sufficient data at the local level to test this claim. Testing the claim requires both program-level COS data and demographic and disability data linked to the COS data to identify appropriate comparisons at entry.

Conclusions for Claim J

There is no conclusion because the claim was not examined.

Claim K

Claim K: COS rating distributions at entry will be related to the disability-related characteristics of the population served by the states.

Rationale for Claim K

IDEA allows states to set their criteria for developmental delay, resulting in differences across states in the percentage of the population receiving early intervention services. It is generally assumed that states serving larger percentages of children in EI are serving more children with milder impairments who would not be found eligible for services in other states. If this assumption is true and the COS rating accurately captures children's functioning at entry, we would expect to see that states serving a lower percentage of children (i.e., providing services only to children who have more severe impairments to their functioning) would have lower COS ratings on average than states serving a higher percentage of children with a broader range of severity in their impairments. This claim assumes that ChildFind activities are comparable from one state to another and that states with lower percentages serve fewer children because of stricter eligibility criteria not because of limited outreach to find families and children who need services. No nationally available data are available to examine what factors might contribute to the percentage of children served in EI in a state in addition to the eligibility criteria.

Evidence for Claim K

To test Claim K, we used national data that states submit annually to OSEP about the children participating in EI and ECSE.

Subclaim K1: National analysis comparing the percentage of young children served by the EI or ECSE to the percentage of children exiting EI or ECSE at age expectations for states using the COS in 2012-13. (National data)

Exhibits 6.5 and 6.6 in Study 4 show the relationships between the percentage of children identified as entering at age-expected levels on the COS for each outcome in 2012–13 and the percentage of children the state EI or ECSE program reported serving in the state.

- The national data for 37 EI programs demonstrated the expected stair-step pattern of increasing percentages of children entering programs functioning at age expectations in states serving a higher percentage of children in the EI program (i.e., if a state serves more children, then the populations served in EI should contain more higher functioning children). Given the small number of states divided across three categories, the mean differences in percentages did not reach the threshold for statistical significance. However, the pattern of the relationships is consistent with the subclaim

- The national data for 29 ECSE programs showed less variation with regard to percentage served. In ECSE, the expected stair-step pattern between percentage served and percentage entering at age-expectation was not found for states serving a moderate or low percentage.

Given mixed findings with support in national EI data but not in ECSE data, the evidence for subclaim K1 was mixed.

Conclusions for Claim K

Data showed a consistent pattern of relationships between the percentage of children entering EI rated as at age-expected levels using the COS and a higher percentage of the state population served. These data support the assertion that the COS distinguishes between those who are more and less impaired in functioning. However, the consistent pattern of relationships was not observed in the smaller set of national ECSE data. Given these mixed findings, support for Claim K was mixed.

Claim L

Validity Claim L: Similar populations of children enter programs each year, so functional levels reflected in COS ratings should remain constant without intervening factors (e.g., new eligibility criteria, rigorous quality assurance, or improvement process implemented).

Rationale for Claim L

If the distribution of children's levels of functioning in the overall group of children served by a specific statewide program is fairly similar from year to year and the COS process accurately reflects children's functioning, then we would expect the overall distribution of COS ratings to be fairly consistent from year to year. Examination of these year-to-year distributions of COS ratings should include only state data with a sufficient annual sample size so that consistency in annual COS ratings distributions would be expected.

Evidence for Claim L

Statewide extant data were examined to test this claim across states that provided multiple years of data. Three states provided three consecutive years of data (either 2008–09, 2009–10, and 2010–11 or 2009–10, 2010–11, and 2011–12), and one state provided two years of data (2009–10 and 2010–11). No specific information was available on major quality assurance initiatives going on in these states during these times or major changes in eligibility requirements or cost structures that might have influenced the distribution of families participating statewide in the EI or ECSE program. However, during that time period, most states did have some quality assurance activities to try to improve the quality of COS data, and the national recession tightened many state budgets. The latter sometimes resulted in staffing changes or policy shifts that could have had an impact on the families served by the programs.

Subclaim L1: Among states with few indicators of problematic COS quality, state distributions of entry COS ratings will not have shifts of more than 3 percentage points from one year to the next. (Extant Data Study)

- There were no year-to-year changes of greater than 3 percentage points for any of the ECSE states (Study 4, Exhibit 6.7).

- This also held true for most of the EI programs. Changes of greater than 3 percentage points between 2008 –09 and 2009–10 were found in one state for 38% of the comparisons. This state is a small, with COS ratings on approximately 500 children, which means that the cell sizes for each of the individual ratings is relatively small and therefore more subject to year-to-year fluctuations. For the larger states, the year-to-year distributions were stable across all the rating points. Eighty percent of the states (four of five) meet the criteria for this claim (Study 4, Exhibit 6.7).

Conclusions for Claim L

Using extant data from four states with medium to large state samples to examine COS distributions at entry across multiple years, all states showed less than a 3 percentage point change on year-to-year percentages in each of the seven COS rating categories. On the basis of these findings, subclaim L1 and claim L were supported.

Claim M

Claim M: Functioning, as reflected in the COS rating, in an outcome area at time 1 is related to functioning in that outcome area at a later point in time.

Rationale for Claim M

Developmental research has demonstrated repeatedly that the best predictor of a child's current development is the child's earlier developmental status. Given that COS ratings reflect the child's level of functioning, we would expect that COS ratings for the same child across time would show a positive relationship to one another.

Evidence for Claim M

Data to examine relationships between COS ratings over time came from extant state data (Study 4) and the smaller research sample from the child assessments study (Study 2). Specific criteria focused on the strength of correlations across time and the percentage of children who displayed unexpectedly large changes over time.

Subclaim M1: COS ratings in the same outcome area at time 1 and time 2 will show moderate correlations (greater than or equal to .5). (Extant Data, Child Assessments Study)

- In a longitudinal sample of the child assessment study (*n* = 70) (Study 2), correlations between entry and exit on the same outcome were .46 for positive social relationships, .47 for knowledge and skills, and .47 for action to meet needs. These are just below the .5 hypothesized correlation set for this subclaim, lending support for subclaim M1.
- In the extant data study (Study 4), data were examined for nine EI programs and nine ECSE programs. Correlations greater than .5 were found for
  - 67% of state EI programs and 100% of state ECSE programs on positive social relationships
    - For EI, state correlations ranged from .38 to .54; for Part B, state correlations ranged from .51 to .66.
  - 33% of state EI programs and 100% of state ECSE programs for knowledge and skills
    - For EI, state correlations ranged from .37 to .51; for Part B, state correlations ranged from .50 to .65.
  - 33% of state EI programs and 100% of state ECSE programs for action to meet needs.
    - For EI, state correlations ranged from .36- to .52; for Part B, state correlations ranged from .50 to .64.

Subclaim M1 was supported by the extant state ECSE data, but support was mixed for the extant state EI program data. Across the three outcomes, this subclaim had the strongest support for positive social relationships, with only three of nine EI programs meeting the subclaim for the other two outcome areas. Taken together, these results support subclaim M1.

Subclaim M2: Less than or equal to 20% of children will show a change of four or more rating points between time 1 and time 2. (Extant Data, Child Assessments Study)

To meet the threshold to support subclaim M2, we would expect most children in the state to maintain or improve their status relative to same-age peers, but few should show large increases or decreases between entry and exit from EI and ECSE. Large changes in status relative to same-age peers between entry and exit from the program are possible but expected to be rare. To meet this expectation, 20% or less of children should have a four-point or more increase or a four-point or more decrease in their COS rating between entry and exit.

- The child assessments study (Study 2) found that three (4%), three (4%), and four (6%) of children had changes of more than four rating points between entry and exit from the program on the three outcomes, respectively. All children with large changes were in a positive direction, consistent with having benefited from an intervention. These percentages are within the threshold identified to support subclaim M2.
- In the extant state data from nine EI and nine ECSE (Study 4), 100% of states had 20% or less children showing changes of four or more points on each of the outcomes. The numbers of instances of both large increases in COS ratings and large decreases in COS ratings were considered. Using simple unweighted state averages, the average

percentages of absolute value of four or more point changes in COS ratings between entry and exit in states was 9%, 10%, and 10%, respectively, for positive social relationships, knowledge and skills, and action to meet needs. Therefore, analyses of extant state data supported subclaim M2, with the majority of the states meeting the subclaim across all outcomes and in all state programs.

Both the child assessments study and the extant data study supported this subclaim; subclaim M2 was supported.

Exhibit 7.10   Summary of Findings for Claim M

Claim M: Functioning, as reflected in the COS rating, in an outcome area at time 1 is related to functioning in that outcome area at a later point in time	

Subclaim M1: COS ratings in the same outcome area at time 1 and time 2 will show moderate correlations (>.5).	Supported
	  In child assessments study data
	Supported
	  In ECSE extant data (100% for all
	  outcomes)
	Mixed
	  In EI program extant data (67%, 33%, 33%)

Subclaim M2: ≤ 20% of children will show a change of more than four rating points between time 1 and time 2.	Supported
	  In child assessments study data
	  In extant data from both EI and ECSE
	  programs.

Conclusions for Claim M

Across the two subclaims, data from the child assessments study and the extant state data study supported claim M (Exhibit 7.10).

Claim N

Claim N: Programs will show differences in the percentage of children making greater than expected.

Rationale for Claim N

COS ratings were developed for accountability purposes. Examining aggregate information about the percentage of children within a given program or state who make greater than expected growth resulting in a change in their developmental trajectories (i.e., Summary Statement 1) is expected to provide important information about which programs or states are having a positive impact on child outcomes. This claim focuses on whether or not the COS can detect aggregate information about the percentage of children who change their developmental trajectories and is sensitive enough to show variation in these percentages across programs or states. Examining these differences can provide one type of information about the effectiveness of the program and is an important intended use of the COS data.

Evidence for Claim N

Evidence to test claim N was drawn from national data and the extant data study (Study 4) and the child assessments study (Study 2). These studies provided information about whether or not changes in trajectories were observable in the data, but they did not have additional information that could confirm whether or not the programs where different percentages of children with trajectory changes were observed were effective.

Subclaim N1: COS ratings can detect a portion of children who change developmental trajectories between time 1 and time 2. (National Data, Child Assessments Study)

- In the child assessments study (Study 2), the COS identified a portion of children who entered below age expectations on the outcome and positively changed their developmental trajectory before exiting the program (also referred to as Summary Statement 1 in state annual reporting to OSEP). The percentage of children entering below age-expected levels and identified as trajectory changers was 58% for positive social relationships, 71% for knowledge and skills, and 63% for taking action to meet needs.

- National data (Study 4) also showed that states were able to report on the percentage of children statewide who entered below age expectations and positively changed their developmental trajectory before exiting the program. All states using the COS provided this information to OSEP as part of annual reports. Exhibits 6.9 and 6.10 (in Study 4) show national data, with the distribution of statewide percentages for Summary Statement 1 from all states using the COS process. These exhibits show percentages for EI and ECSE on the action to meet needs outcome as an example. Exhibit 6.11 summarizes the range of values observed across states on all three outcomes.
  - For EI, for the 37 states in the extant data study, the percentages pf children who changed trajectories ranged from 24% to 91%, from 37% to 95%, and from 48% to 92% for positive social relationships, knowledge and skills, and actions to meet needs, respectively.
  - For ECSE, for the 29 states in the extant data study the percentages pf children who changed trajectories ranged from 46% to 96%, from 64% to 98%, and from 57% to 95% for positive social relationships, knowledge and skills, and actions to meet needs, respectively.

These data show that COS ratings can be used to identify a group of children who enter below age expectations and positively change their developmental trajectories before exiting the program. Percentages can be computed for Summary Statement 1 and show state-to-state variation. These data support subclaim N1.

Subclaim N2: Summary Statement 1 values reflecting positive changes in developmental trajectories among children who enter below age expectations will show a range of values across local programs. (Extant Data Study)

This subclaim is supported if local or regional programs within states show variance in the percentage of children making greater than expected growth (i.e., changes in their developmental

trajectories as evidenced in Summary Statement 1). To examine this claim, the range of Summary Statement 1 (greater than expected growth) values was computed across local or regional programs. To be included in the analysis, local or regional programs were required to have at least 30 children exiting. Data were included from seven EI state programs and eight ECSE Programs that provided data on local or regional programs.

Exhibit 6-12 (in Study 4) shows the range of values found across local or regional programs. These data show that for all states with local or regional program data in the extant data study, there was local/regional variation across both EI and ECSE programs in percentages of children classified as making greater than expected growth (Summary Statement 1). Variation was observed for all states and for all three outcomes.

These data across states' local programs support subclaim N2.

Exhibit 7.11   Summary of Findings for Claim N

Claim N: The rating structure of the COS is sensitive to both improvements in and maintenance of developmental trajectories that occur in effective programs	
Subclaim N1: COS ratings can detect a portion of children who change developmental trajectories between time 1 and time 2.	Supported
Subclaim N2: Summary Statement 1 values reflecting changes in developmental trajectories among children who enter below age expectations will show a range of values across local programs.	Supported

Conclusions for Claim N

Both criteria for claim N were supported, indicating the COS ratings can be used to identify those who change developmental trajectories and in aggregate show variation necessary to use this information for program improvement. Therefore, the evidence supports claim N (Exhibit 7.11).

Claim O

Claim O: Data produced by the COS are sufficiently precise to enable states to track the overall status of their EI or ECSE system with the summary statements and monitor change toward targets on those summary statements.

Rationale for Claim O

Many states are aggregating COS data to produce progress categories and summary statements to meet federal reporting requirements (see Background, Section 1, for information about progress categories and summary statements). As part of OSEP's established approach for accountability, states set targets for the summary statements and monitor changes in summary statements relative to those targets. For COS data to be valid for this purpose, the summary statements produced from COS data must remain relatively stable across years under similar

conditions and be sensitive enough to demonstrate increases and decreases in response to statewide improvements and challenges.

Evidence for Claim O

Investigating the sensitivity of summary statements to statewide EI and ECSE improvements and challenges was beyond the scope of the set of ENHANCE studies that were undertaken. However, national data submitted for federal reporting by states using the COS process were analyzed to investigate whether summary statements showed some patterns of stability or incremental change across years or showed extreme fluctuations from year to year.

Subclaim O1: Few states will show patterns with erratic shifts in summary statements from year to year.

National summary statements submitted for federal reporting by state EI and ECSE programs were used to test subclaim O1. Only data for states using the COS for accountability between 2009–10[35] and 2012–13 and meeting minimum quality criteria to be appropriate for the national analysis were analyzed. This resulted in use of data from 20 state EI and 18 state ECSE programs.

Exhibits 6.13–6.15 (in Study 4) are plots of the changes in Summary Statement 1 percentages for each of the three outcomes in state EI programs across years. Of the 20 state EI programs, few made big shifts in Summary Statement 1 percentages. Most states showed fairly stable or slightly incremental trajectories.

Exhibits 6.16–6.18 (in Study 4) are plots of the changes in Summary Statement 1 percentages for each of the three outcomes in state ECSE programs across years. Of the 18 state ECSE programs, few made big shifts in Summary Statement 1 percentages. Most states showed fairly stable or slightly incremental trajectories.

In summary, for both state EI and ECSE programs with Summary Statement 1, this subclaim was supported.

Exhibits 6.19–6.21 plot the changes in Summary Statement 2 percentages for each of the three outcomes in state EI programs across years. Of the 20 state EI programs, few had big shifts in Summary Statement 2 percentages. Most states showed fairly stable or slightly incremental trajectories.

Exhibits 6.22–6.24 plot the changes in Summary Statement 2 percentages for each of the three outcomes in state ECSE programs across years.

- Of the 18 state ECSE programs, few had big shifts in Summary Statement 2 percentages for the knowledge and skills outcome. Most states showed fairly stable or slightly incremental trajectories on this outcome.

---

[35] Most states had data from 2008-09, however often sample sizes shifted in states during those early years as states expanded outcomes data collection statewide and address major training issues influencing data quality. So, data shown reflect 2009-2013.

- For positive social relationships and action to meet needs, about one-third of states had a moderate amount of fluctuation in Summary Statement 2 percentages. Most of the states with larger changes had fluctuations in these outcomes in their 2010–11 data.

In summary, for state EI programs with Summary Statement 2 the evidence of stability was strong, while for state ECSE programs stability was more moderate.

Visual inspection of stability supported subclaim O1 in 100% of the exhibits for EI across both summary statements and all three outcomes. For ECSE, 100% of the exhibits for Summary Statement 1 were supported and 33% of the exhibits for Summary Statement 2 were supported. Taken together, the preponderance of evidence from the extant state data about stability supports subclaim O1.

Conclusions for Claim O

Visual inspection confirmed considerable stability or mild incremental growth/drops in Summary Statement 1 percentages by state programs across years. This finding was demonstrated strongly in state EI programs and in four of six exhibits showing stability trends for state ECSE programs. Therefore, claim O was supported.

Claim P

Claim P: Providers will report minimal negative consequences in practice as a result of implementing the COS process.

Rationale for Claim P

The COS process is based on best practices in early childhood assessment because the assessment of the three broad functional outcomes is expected to be based on information about children's functioning across settings and situations (Division for Early Childhood, 2007; National Association for the Education of Young Children, 2009). As described in the introduction in Section 1, stakeholders involved in formulating the OSEP child outcomes encouraged the use of outcomes that were *sufficiently global* to be relevant to children with all types and severities of disabilities. Global outcomes also would enable states to map the OSEP outcomes to the varied early learning standards that exist in many states. Stakeholders strongly supported making the child outcomes *functional* (i.e., outcomes that have meaning in the child's everyday life). A clear message was that the outcomes should *not be built around traditional child developmental domains* (e.g., cognition, motor skills) because recommended practice is to think about children from a transdisciplinary perspective that focuses more on how children integrate skills across these domains in functional ways(McWilliam, 2004). The stakeholders felt strongly that a domains-based approach to outcomes would reinforce an outdated intervention model of individual therapists addressing skills limited to those covered in their professional disciplines.

Thus, because the COS process requires use of multiple sources of information about children's functional skills and reflects recommended practice for EI and ECSE assessment and practice(Division for Early Childhood, 2007, 2014; National Research Council, 2008), we hypothesized that providers would not view the COS process as having negative impacts on their practice.

Evidence for Claim P

The evidence for this claim came from the online provider survey (Study 1), namely the items in which EI and ECSE providers rated their perceptions of the potential negative impacts of the COS process on their practice.

Subclaim P1: At least 85% of providers will indicate not at all true that the COS process has significant negative consequences of the COS with regard to"The ratings were selected to make the program look good." (Provider Study)

A majority of providers participating in the online survey (88%) reported not at all true for this statement (Study 1, Exhibit 3.10). Therefore, this subclaim was supported.

Subclaim P2. At least 85% of providers will indicate not at all true that the COS process has significant negative consequences of the COS with regard to "The COS process takes time away for other important activities." (Provider Study)

A majority of providers participating in the online survey (88%) reported not at all true for this statement (Study 1, Exhibit 3.15). Therefore, this subclaim was supported.

Subclaim P3: At least 85% of providers will indicate not at all true that the COS process has significant negative consequences of the COS with regard to "The COS process negatively impacts the assessment process." (Provider Study)

A majority of providers participating in the online survey (75%) reported not at all true for this statement (Study 1, Exhibit 3.15). Therefore, this subclaim was mixed.

Subclaim P4: At least 85% of providers will indicate not at all true that the COS process has significant negative consequences of the COS with regard to "The COS process has negative impacts on my relationships with families." (Provider Study)

A majority of providers participating in the online survey (87%) reported not at all true for this statement (Study 1, Exhibit 3.15). Therefore, this subclaim was supported.

Subclaim P5: At least 85% of providers will indicate not at all true that the COS process has significant negative consequences of the COS with regard to "The COS process leads to poorer quality IFSP of IEP outcomes." (Provider Study)

A majority of providers participating in the online survey (90%) reported not at all true for this statement (Study 1, Exhibit 3.15). Therefore, this subclaim was supported.

Subclaim P6: At least 85% of providers will indicate for all four negative impact items above (P2-P5) that they were not at all true. (Provider Study)

A majority of providers (77%) indicated no negative impacts for all four of these impact items (Study 1, Exhibit 3.15). Therefore, the evidence for this this subclaim was mixed.

Subclaim P7: At least 85% of providers will <u>not</u> indicate that "overall the COS process has negative or very negative impact on my work with children and families." (Provider Study)

A majority of providers (93%) did not report negative or very negative as the overall impact of the COS on their work. Therefore, this subclaim was supported.

Exhibit 7.12   Summary of Findings for Claim P

Claim P: Providers will report positive changes in practice as a result of implementing the COS process and minimal negative consequences	
At least 85% of providers will….	
Subclaim P1: ... indicate not at all true that the COS process has significant negative consequences of the COS with regard to "The ratings were selected to make the program look good."	Supported (88%)
Subclaim P2: ... indicate not at all true that the COS process has significant negative consequences of the COS with regard to "The COS process takes time away for other important activities."	Supported (88%)
Subclaim P3: ... indicate not at all true that the COS process has significant negative consequences of the COS with regard to "The COS process negatively impacts the assessment process."	Mixed (75%)
Subclaim P4: ... indicate not at all true that the COS process has significant negative consequences of the COS with regard to "The COS process has negative impacts on my relationships with families."	Supported (87%)
Subclaim P5: ... indicate not at all true that the COS process has significant negative consequences of the COS with regard to "The COS process leads to poorer quality IFSP of IEP outcomes."	Supported (90%)
Subclaim P6: At least 85% of providers will indicate for all four negative impact items above (P1P2-P5) that they were not at all true.	Mixed (77%)
Subclaim P7: ... **not** indicate that "overall the COS process has negative or very negative impact on my work with children and families."	Supported (93%)

Conclusions for Claim P

Across seven criteria for Claim P, all were either supported (five subclaims) or mixed (two subclaims) (Exhibit 7.12). Taken together, Claim P was supported by the evidence from the online provider survey data.

Overview Summary of Validity Claims

The ENHANCE project provided an opportunity to identify and test a series of 16 validity claims using data from four studies. Differential strengths and weaknesses across these studies enabled project staff to gather a wide range of evidence about the validity claims and consider the extent of support for each of them. Exhibit 7.13 provides a summary of the conclusions regarding the 16 validity claims.

The four studies provided considerable evidence for the validity of the COS ratings for purposes of local and state accountability for early intervention and early childhood special education programs. Most of the validity claims were tested empirically (13 of 16; 81% of the claims). Of the 13 claims tested across the four studies, 11 claims were supported and 2 were partially supported.

Exhibit 7.13   Summary of Conclusions on Validity Claims About the COS Process

	Claim	Summary
A	The three outcome areas reflect important outcomes for accountability and program improvement.	Supported through process surrounding identification of outcomes. Not empirically tested in this project.
B	The skills and behaviors described in the COS guidance materials and training resources are sufficient to enable providers to accurately describe the content of the three OSEP outcomes.	Supported.
C	There is variability in children's functioning in the three outcome areas, and that variability is reflected in the COS ratings.	Supported.
D	There are developmental sequences within each outcome that provide the internal structure of the COS ratings.	Supported via literature. Not empirically tested in this project.
E	Providers of EI and ECSE services can be trained to understand and correctly apply knowledge of child development and the COS rating criteria such that a child's team will assign an accurate rating.	Supported.
F	Functioning in one outcome area is related to functioning in another outcome area.	Supported.
G	COS ratings in the corresponding outcomes are moderately related with the social emotional (outcome 1), cognitive (outcome 2), communication (outcome 2), and adaptive (outcome 3) domain scores of assessment tools.	Mixed.
H	COS ratings will be related to the type and severity of the child's disability.	Supported.
I	COS ratings will not be related to the composition of the team, the particular assessment tool used, or child characteristics such as gender or race/ethnicity among children with the same types and severities of disabilities.	Supported.
J	COS entry scores will be similar for programs and regions serving similar populations.	Not empirically tested on this project.

		Claim	Summary
K		COS rating distributions at entry will be related to the disability-related characteristics of the population served by the states.	Mixed.
L		Similar populations of children enter programs each year, so functional levels reflected in COS ratings should remain constant without intervening factors (e.g., new eligibility criteria, rigorous quality assurance, or improvement process implemented).	Supported.
M		Functioning, as reflected in the COS rating, in an outcome area at time 1 is related to functioning in that outcome area at a later point in time.	Supported.
N		The rating structure of the COS is sensitive to both improvements in and maintenance of developmental trajectories that occur in effective programs (i.e., COS ratings differentiate effective from ineffective programs).	Supported.
O		Data produced by the COS are sufficiently precise to enable states to track the overall status of their EI or ECSE system with the summary statements and monitor change toward targets on those summary statements.	Supported.
P		Providers will report minimal negative consequences in practice as a result of implementing the COS process.	Supported.

Section 8 – Summary and Conclusions

The ENHANCE project was designed to address three objectives:

1. Conduct a program of research to examine the validity of ratings generated by the COS process and identify conditions that lessen validity.

2. Revise the COS form and supporting materials based on study findings.

3. Identify of a series of validity analyses that can feasibly be conducted in states to enable each state to examine the validity of its COSF data on an ongoing basis.

This final section of the report summarizes the project's accomplishments relative to each of these objectives.

Objective 1: Conduct a Program of Research

Summary of Findings

The project team carried out four studies to examine the validity of the Child Outcomes Summary process. Those studies involved surveying providers who were implementing the COS process, comparing COS ratings with the results of two child assessments, examining COS team meetings through the use of videotapes, and analyzing data from state EI and ECSE programs using the COS.

- The findings from these studies were used to test 13 validity claims.
- Support was found for 11 of the claims, and the evidence was mixed for the other two.
- Two additional claims were supported with evidence collected outside the study
- One claim could not be tested because no data were collected.
- Overall, support was found for 13 of the 15 claims examined.

Key study findings regarding the validity of the COS addressed providers' understanding of the content of the outcomes, the patterns in the ratings, the relationship of the ratings to other assessment tools, expected patterns for subgroups, and the lack of negative consequences of implementing the COS process. Together, these findings provide support for the validity of the COS.

Specifically, the majority of providers reported that they understood the content of the three child outcomes, and this was substantiated by the videos. When they were observed in team meetings, providers generally displayed an accurate understanding of the content of the outcomes and correctly associated skills with the outcomes in the documentation and on the videos. Providers were observed on the videos correctly identifying the ages at which children are expected to acquire certain skills, which is a prerequisite for being able to link the child's skill level to a COS rating. Providers also reported in the survey that they understood age-expected functioning in the three outcome areas. When discussion during the video of a team

231

meeting was sufficient for an external coder to also determine a rating, the coder generally assigned a rating within one point of the team's rating.

Several analyses across studies found that the ratings produced by the COS process spanned the full range of ratings and percentages with extreme values of ratings that were consistent with what is known about the populations served in EI and ECSE. Relatively few children had ratings of 1, and only a small percentage of children were rated as having age-expected functioning in all three outcome areas at entry to EI or ECSE. The full range of ratings (i.e., 1 to 7) was found in each of the state data sets.

Relationships found between COS ratings and assessment scores and in the state COS data were consistent with what would be expected if the COS ratings accurately captured children's functional levels in the three outcomes. Across multiple studies, strong relationships were found between the three outcome areas at entry and exit. Moderately strong relationships also were found between entry and exit for each outcome. Various domains on the BDI-2 and Vineland-II were moderately related to the relevant outcomes, with the BDI showing stronger relationships to COS data than the Vineland.

The COS ratings differed as expected for different subgroups of children. For instance, comparisons showed the expected patterns for children with low or high functioning as represented by ABILITIES Index scores, for those who entered at age expectations on all three outcomes relative to those who did not, and for those with specific patterns of functioning expected based on their primary disability category. An analysis of data for all states using the COS showed the expected relationships between state-level COS data and the percentage of children in the state served by EI or ECSE, with the relationship being strong for the EI programs. Within states, there was variation across local programs in the percentage of children deemed to have made greater than expected growth. The outcomes data reported to OSEP by states that use the COS have been very stable from year to year for most of the states. Finally, few providers reported any negative consequences associated with collecting information about young children with the COS process.

Not all the findings conclusively suggested that all data produced from the COS process are error free and valid, and some of the evidence was weaker than expected. The study uncovered several issues related to implementation of the COS process and the background understanding of some providers on COS decision-making teams. The dialogue observed in videos of the team decision-making process showed that teams did not consistently discuss the child's functioning in sufficient depth to create a full picture of the child's skills and behaviors or with regard to the breadth of each of the three outcomes. This limited dialogue was related to the brevity of these meetings; some teams did not spend enough time together to pursue the level of discussion that would have been considered ideal or even adequate for thoughtful decision making. As outsiders to the team process whose access to the team's interactions was limited to the video, we could not know whether the brief discussion observed was the sum total of the information shared about the child's functioning or whether it was the culmination of multiple communications that

had occurred before the team meeting. Although many of the discussions did not reflect our vision of a high-quality team process, we may have been observing needed and appropriate efficiencies by providers with very limited time. Much remains to be understood about how much of the information about children's functioning is shared among team members in other ways or via shorthand references that are not as apparent or understandable to outside observers. Alternatively, providers may need to exchange very little information to pinpoint where the child is on the COS scale.

Although providers were confident about their own knowledge about the COS process (e.g., understanding of the three outcomes, knowledge of skills expected at different ages, and understanding of the rating criteria for the COS), they were slightly less confident in their fellow team members' knowledge. Providers reported difficulty both understanding and explaining the purpose of the COS and limited knowledge of how the data are being used. Some of these findings are not surprising because 75% of the providers reported having received four hours or less of COS training. The national technical assistance center supporting child outcomes measurement recommends that at least one full day of training is needed to understand and allow for guided practice time with the COS process. Providers also noted that only half their programs ever provided feedback to those involved in the COS process, and only slightly more than one-third of the providers were in programs where someone checked their COS forms to ensure the accuracy of the ratings. Providers' self-reported lack of understanding of why the data are being collected is problematic in that they might perceive the COS process to be of limited value especially because it was not designed to provide any additional information to guide practice. Engaging in what is perceived to be a meaningless activity provides limited motivation to invest the time needed for accurate rating decisions. Most likely these implementation issues negatively affected the quality of the COS data and weakened some of the evidence for the validity claims to an unknown extent.

The findings from the child assessments study merit additional discussion because demonstrating a relationship with other assessment tools is a common component of a validity argument. We expected from the outset that this claim would be problematic because there are no existing assessment tools that address the content of the three outcomes. There was no expectation that a domain score on a domain-based tool would be an adequate measure for any of the outcomes. In fact, the COS, which is not an assessment tool but a summary process, was developed to address the limitations of existing tools as measures of the three outcomes. Also, computing correlations between COS rating and assessment scores presents statistical challenges. The COS has two ratings, 6 and 7, to capture the diversity of functioning within the age-expected range, whereas the BDI-2 and the Vineland-II have a full range of scores at and above age-expected functioning. The limited range on the upper end of the COS ratings would be expected to attenuate the correlations with these assessment tools. These issues notwithstanding, young children with disabilities span a broad spectrum of skills and abilities, and children with severe impairments would be expected to score low on any measure of their skills or functioning and children with mild impairments would be expected to score much higher. With limited data

upon which to base a prediction, we predicted that there would be moderate relationships between the COS ratings and the BDI-2 and the Vineland-II.

Child assessment study analyses found moderate relationships for COS ratings with the BDI-2 (.41 to .46) and low to moderate relationships with the Vineland-II (.28 to .50) across the three outcomes. Ironically and unexpectedly, the highest correlation for each assessment was between the third outcome, takes action to meet needs, and the motor domain on each tool (.46 and .50, respectively). We expected this relationship to be one of the lowest because the assessments address motor skills and the outcome and the COS addresses mobility, which, with assistive technology, can be very different from motor skills. Although the correlations between the COS ratings and the two assessment tools were in the low to moderate range, this is probably what is to be expected given that the COS ratings and the assessment tools measure different aspects of a child's development.

A more relevant analysis, considering that the COS was designed for accountability reporting, was the distribution of progress categories and percentages for the summary statements produced by the COS relative to what was produced by the other two tools. These are the metrics that states are required to report to OSEP, so an important question is whether these different approaches produce a similar aggregated picture for federal reporting purposes. This is not a theoretical question since some states use the BDI-2 for child outcomes reporting. The answer, as presented in Section 4, is that they do not. The finding is difficult to interpret with regard to the validity of the COS for several reasons. Data are based on the small sample size available for the analysis ($n = 70$), and no clear guidance exists about how BDI-2 and Vineland-II scores should be converted to progress categories.[36] Because the BDI and Vineland were not designed for this purpose, many critical decisions are needed with regard to how to map domains to outcomes and to operationally define the threshold for age-expected functioning, changes in developmental trajectories, and a child's acquisition of skills over time. Different decision rules produce different distributions of progress categories and therefore different percentages of children for the summary statements, even with the same sample of children on the same assessment tool. Although these different progress category distributions and summary statement frequencies obviously cannot all be accurate, there is no way to determine whether the COS or the BDI-II produces the more accurate information.[37]

---

[36] In fact, there is no uniform approach for this conversion across states currently using the BDI-2 for accountability reporting, and differences have been observed when different assumptions have been applied.

[37] Before introduction of the three child outcomes, children in EI and ECSE had never been measured using the same assessment tool or approach. Therefore, the appropriate distribution for this population, even using domain-based assessment tools, is unknown.

Lessons and Implications for Future Research

Validity of a measurement strategy is best examined under ideal implementation. If the approach cannot yield valid data for the intended purpose under ideal conditions, it certainly will not produce valid data under less than ideal conditions. When an approach is studied under less than ideal implementation, as the COS process was in this project, uncertainty will exist about the extent to which limited implementation negatively affected the findings. The design of the project called for recruiting EI and ECSE programs in states that were recognized by their state agency as good implementers of the COS process. Unfortunately, site recruitment occurred at the height of the recession, when programs were struggling with budget and staffing issues and many were not willing to be part of a research project. As the difficulty in recruiting sites considered to be good implementers became more apparent and site recruitment extended over many months, we broadened the criteria to include any sites recommended by state agency staff that would participate in the project. Casting a broader net for site recruitment meant that some of the sites were doing a better job of implementing the COS process than others.

We also encountered challenges in recruiting families within those sites, including some sites that ultimately recruited none or only one family or participated in data collection for some ENHANCE studies but not others. Recruiting families into the assessment study or the video study was somewhat burdensome for program staff in that they were required to explain the study and its processes to families new to the program and encourage those families to set aside time for study assessments in addition to program assessments that were already planned with the child. In the case of the video study, a number of consents had to be collected from all team members in addition to the family and the COS meeting had to be videoed. Staff turnover in the sites was another barrier to recruiting families and submitting timely data to the project. The ultimate consequence of difficulties and delays in timing of recruiting families was a smaller sample size for the child assessments study than we had wanted, with some of these families being in programs that were less than ideal implementers of the COS process. One of the consequences of the smaller than expected child sample was that some of the planned multivariate analyses and investigation of differences between EI and ECSE programs could not be conducted.

Despite the challenges with both site and family recruitment and the less than ideal implementation as documented by both the provider survey and the team decision-making study, the findings do support the validity of the COS for accountability reporting, the purpose for which it was designed. One interpretation of the findings is that the COS is a robust process that can yield valid data for accountability purposes even when providers are not thoroughly trained or spending much time discussing the child's functioning in each of the outcome areas. Although we are reassured by the COS data, we would nevertheless strongly advocate for more training of COS team members, better implementation, and more systematic checks on the quality of the data based on what was learned through the studies.

The research had other limitations in addition to the implementation challenges. It was conducted by the researchers who developed the COS process. Before and during the study, the research team also was part of a national center providing states with technical assistance on measuring child outcomes and on implementing the COS process. We worked to be unbiased in our data collection, analysis, and interpretation, but each of us has invested many years in the COS process, and our closeness to the focus of the study needs to be acknowledged.

A second limitation was that all the participants were aware that they were being studied. The providers had to recruit the families for the study, so they knew which COS ratings we would be comparing with the assessment scores (which were independently obtained by assessors hired and trained by the project). More important, each of the teams being videoed knew about the video recording and that their team process would be studied. Given the diversity of meetings seen in the video, it is a safe assumption that not all the teams adjusted their behaviors for the sake of the camera, but some of what was seen might have been influenced by the filming in unknown ways.

Additional research on the validity of the COS process under ideal and systematically varied conditions would be helpful to address such important implementation questions as

- How much and what kinds of training and experience do providers need to accurately determine COS ratings?
- How much discussion is needed for a team to produce an accurate rating?
- What is the effect of periodic monitoring and feedback on the quality of the data?
- Do providers need refresher courses on the COS process and, if so, how often?

We need to know more about the conditions under which the validity of the data is enhanced and lessen. This findings from this study established that the COS process can produce valid data. Given that validity is not characteristic of the COS process but of the data and use of the data and that there is substantial variation in implementation, there remains a need to identify the critical factors that affect the validity of the data.

The importance of the COS process to national data on child outcomes for children receiving EI and ECSE has continued to grow in the years since this study was initiated. Including the data submitted in February 2015, states have now reported six years of national child outcomes data to OSEP. In 2012–13, the most recent year for which data are available, 75% (42 of 56) of the states and territories used the COS as their measurement strategy for child outcomes for EI. For ECSE, the figure was 63% (37 of 59). These data are being used nationally; the U.S. Department of Education includes the state-reported data on child outcomes in the President's annual budget requests for these programs. The validity of the COS data gained additional significance when OSEP recently introduced a new requirement for states to produce a State Systemic Improvement Plan (SSIP). To develop the SSIP, states are required to analyze their data (including child outcomes data), to identify a measurable result, and to design and implement strategies to improve the result. States also are to use data to monitor the success of their efforts. Many states already have identified a specific child outcome (e.g., positive social

relationships) as their focus. States will be tracking their child outcomes data with additional scrutiny in the years ahead since changes in the outcomes data will serve as the marker of the success of their improvement plan.

Starting in 2016, OSEP will use the child outcomes data for a new purpose. As part of Results Driven Accountability (RDA) and reflecting a shift from monitoring for compliance to monitoring for results, states for the first time will be monitored by OSEP on their child outcomes data. The monitoring process identifies the lowest performing states as "needing improvement" and requires certain actions on the part of the state. With this new use, the child outcomes data will become high stakes for states in a way that they have not been before. Using the data for monitoring introduces a new condition with regard to the validity because states will be under more pressure to show good outcomes. Providers may learn about this use and see greater value in collecting quality COS information. However, if this pressure is communicated to providers, it also could negatively affect the accuracy of the rating process. Safeguards to monitor the implementation of the COS process and checks on the validity of the data will become even more essential when the data are used for monitoring.

Objective 2: Revise the COS Form and Supporting Materials

During the ENHANCE project, the research team continued in its role as technical assistance provider to states for child outcomes measurement. We used what we were learning from ENHANCE and from our ongoing interactions with states to improve the materials available to states to support the COS process. The entire set of COS resources is on the child outcomes page of the Early Childhood Technical Assistance Center's (ECTA) website (http://ectacenter.org/eco/pages/outcomes.asp#COSFormandInstructions) and many are in Appendix A. We did not learn anything through the ENHANCE study that suggested a need to revise the COS form itself. There are two versions of the form on the website, and many individual states have developed their own versions, which we have encouraged.[38] We did learn much, however, about the need for materials to support states in implementation of the COS process.

Some of the newest materials were informed by the findings of the ENHANCE project. A self-guided online learning module on the COS process is under development, with a scheduled release time of early summer 2015. The module consists of six narrated learning sessions, resources, supplemental activities, and assessments. It covers many of the same topics that are in face-to-face COS trainings, such as the background and contents of the three outcomes, why the data are being collected, what teams need to know to determine a COS rating, the definitions for each of the rating points, and how to document the team's decision. The hope is that the module will address some of the issues uncovered in the provider survey, such as a lack of understanding

---

[38]   States are encouraged to incorporate the form into other existing state materials and format it as needed for their systems. However, core features of the questions and ratings are expected to remain consistent.

of why the data are being collected and the minimal training that providers are receiving. All who interact with the module will hear the same material, which should improve the consistency of training and thereby improve the overall quality of providers' training. We expect that those who complete the module will still need supervised practice in using the COS process in team decision-making situations, but they will go to those practice sessions with stronger foundational knowledge than many providers currently are.

A second product under development is the COS competency check. This is an assessment consisting of two levels, a screener and a more in depth assessment. The screener is a set of multiple-choice questions about the outcomes and the 7 COS rating points. It is designed to screen out individuals with minimal knowledge of the COS process who would almost certainly fail the more in-depth assessment. Those who cannot pass the screener will be instructed to pursue additional training and then retake the screener. The level-two assessment requires the provider to apply knowledge of the COS process. It includes a case study of a young child with detailed assessment information and a series of questions about the appropriate ratings given the information provided. Much of the development of the competency check is completed, and pilot data have been collected; the assessment is being revised on the basis of the information from the pilot testing. The competency check is scheduled to be released in fall 2015. It will provide local program directors, as well as state agency staff if they choose to see the scores, with information on which providers need additional support in providing valid COS ratings. Both the screener and assessment were initiated during the ENHANCE project and have received additional support from ECTA and the Center for IDEA Early Childhood Data Systems, an OSEP-funded project focused on improving data quality.

A third product under development is an observational checklist of quality indicators of the COS process. This work is being led by an early intervention director in consultation with the research team and ECTA staff. It consists of a checklist of the elements of a quality COS process, along with descriptions of each element and a set of illustrative video clips. The coding procedures developed for the video, study as well as what was learned from observing teams conducting COS meetings, contributed to the content of this checklist.

Numerous other resources to support professional development related to the COS process have been developed, and previously existing materials have been revised over the course of the ENHANCE project, based in part on what was learned about professional development needs in the ENHANCE research sites. These materials are on the ECTA website at http://ectacenter.org/eco/pages/training_resources.asp . Included in this collection are narrated presentations, presentation slides, training activities, quizzes, and refresher presentations. These materials were designed to be used by state, regional, and local trainers who are teaching providers about the COS process.

Several states also have used surveys to study COS implementation, and some have reviewed the ENHANCE provider survey as they identified survey questions. One state's experience with its implementation survey was presented at the 2011 Measuring Child and Family Outcomes

Conference. In the presentation, "Do You See What I See? Assuring Fidelity of COS Implementation," Wisconsin representatives shared how the state is using a fidelity self-assessment to work with programs to help them reflect and improve on COS implementation. The presentation is at http://ecoutcomes.fpg.unc.edu/resources/do-you-see-what-i-see-assuring-fidelity-cos-implementation .

Objective 3: Identify Validity Analyses for States

Ensuring the validity of the COS data requires several different kinds of efforts at the state level. One is the provision of high-quality and sufficient professional development so that providers involved in making rating decisions are thoroughly familiar with the key concepts. As discussed, numerous materials to support preparing providers in participating meaningfully in the COS process were developed or revised on the basis of what we learned through ENHANCE. A second but very different kind of activity that supports improved validity is the analysis of the COS data.

States have analyzed child outcomes data to examine the quality of the data, to use the data for program improvement, or both. Although analyzing the data for program improvement is not carried out as a check on validity, our experience in working with states has shown that use of the data uncovers quality issues and the state's desire to use the data provides a strong impetus to identify and address data quality issues. The research team has worked with other technical assistance providers to develop tools to assist states in analyzing their child outcomes data. We have developed tools for states to examine the validity of their data by looking for expected patterns in the data (i.e., a generic set of patterns that form a validity argument). We also have consulted with states on the development of such tools. Finally, we have given numerous presentations and conducted workshops with and for state staff on data analysis and interpretation over the course of the ENHANCE project. A list of ENHANCE dissemination activities is in Appendix N. These tools and presentations are summarized in Exhibit 8.1.

Exhibit 8.1   Materials Developed to Support States in Analyzing COS Data

Support with Pattern Checking

Checking Outcome Data for Quality: Looking for Patterns Description: This document describes strategies for using data analysis to improve the quality of state data by looking for patterns that indicate potential issues for further investigation. The pattern checking table was revised in July 2012 based in part on information learned through ENHANCE. Location: http://ectacenter.org/eco/assets/pdfs/Pattern_Checking_Table.pdf

Patterns in Child Outcomes Summary Data: Analytic Approaches and Early Findings from the ENHANCE Project Description: Presented at the 2011 Measuring Child and Family Outcomes Conference, this was an update on the ENHANCE project. Presenters shared preliminary findings from the analysis of state data and techniques being used for interpreting patterns to understand the validity of the data. Materials were provided to support states in analyzing the quality, consistency, and meaning of their own COS data. Location: http://ecoutcomes.fpg.unc.edu/resources/reporting-local-child-outcomes-data-facilitated-discussion-about-challenges-and

Support with Analyzing and Using Child Outcomes Data for Program Improvement

Guidance Table for Analyzing Child Outcomes Data for Program Improvement Description: This tool, released in September 2013, was designed to help identify key issues, questions, and approaches for analyzing and interpreting data on outcomes for young children with disabilities. It outlines a series of steps related to defining analysis questions, clarifying expectations, analyzing data, testing inferences, and conducting data-based program improvement planning. It also includes examples of questions, approaches, and sample figures to consider. See also presentations and materials on this topic from the Improving Data, Improving Outcomes conference, including one state's approach to using the guidance table. Location: http://ectacenter.org/eco/assets/pdfs/AnalyzingChildOutcomesData-GuidanceTable.pdf

Where the Rubber Hits the Road: Tools and Strategies for Using Child Outcomes Data for Program Improvement Description: Presented initially at the 2013 Improving Data, Improving Outcomes Conference and then again as a national webinar because of its popularity, this presentation describes how programs can use outcomes data to change systems and practices. Presenters reviewed key tools designed by the Early Childhood Outcomes (ECO) Center to help state and local programs analyze, interpret, and use outcomes data. One state shared how its state technical assistance has addressed using the child outcomes data to help local programs improve the quality of their data and change practices. ENHANCE staff helped this state design a series of tables to decide what analysis to do and to share information with local programs about the quality of their data. Location: http://unc-fpg-cdi.adobeconnect.com/rubberroadencore/ (presentation)   http://www.ectacenter.org/eco/assets/docs/Kasprzak-ABC.doc (handout)

Identifying Meaningful Differences Between State and Local Summary Statement Values Description: Presented at the 2013 Improving Data, Improving Outcomes Conference, this session introduced states to the concept of using confidence intervals with their two summary statements for the child outcomes data. The use of a confidence interval enables the state to identify differences between the program and the state numbers that are and are not statistically significant. In this session, a tool for computing these confidence intervals was presented. Three states described their use of confidence intervals for comparing state and local summary statement data. Location: http://ecoutcomes.fpg.unc.edu/resources/identifying-meaningful-differences-between-state-and-local-summary-statement-values

SSIP Child Outcomes Subgroup Analysis Template

Description: This document provides states with table templates for subgroup analyses that have proven useful in understanding predictors of child outcomes. These are suggestions and need to be tailored to fit the appropriate subgroups for the state. These table templates were derived from tables developed for the ENHANCE extant state data study and were made widely available in 2013.

Location: http://ectacenter.org/eco/assets/docs/subgroupdataanalysistemplate.docx

---

Data Drill Down: Supporting Local Programs in Realizing the Possibilities for Using Data

Description: How can local programs implement a process for drilling down into their data to ensure data quality and program quality? In this session, presented at the 2012 Measuring and Improving Child and Family Outcomes Conference, one state shared its Data Drilldown Guide and training developed for supporting local programs in looking at child outcomes data and planning for improvement. The Early Childhood Outcomes Center (ECO) shared a national resource with suggested drill down questions developed for child and family outcomes.

Location: http://ecoutcomes.fpg.unc.edu/resources/data-drill-down-supporting-local-programs-realizing-possibilities-using-data

---

Local Contributing Factor Tool

Description: This document provides the types of questions a local team would consider in identifying factors affecting performance. It was designed to assist local programs in collecting valid and reliable data to determine contributing factors impacting performance on State Performance Plan (SPP) indicators. The latest addition to the existing tool, released in December 2012, is a section of drill-down questions on child outcomes. A video, Local Contributing Factor Tool- New Sections for Child Outcomes (C3/B7), provides an introduction to the tool and the new sections for child outcomes.

Location: http://ectacenter.org/eco/assets/docs/ECO-C3-B7-LCFT.docx

---

Creating a Culture of Using Data at the Local Level

Description: In this session, presented at the 2012 Measuring and Improving Child and Family Outcomes Conference, Part C staff from one state described use of statewide analysis of patterns in child outcomes data collected using the Child Outcomes Summary process. They also described the process of building the capacity of their local programs to understand and use data.

Location: http://ecoutcomes.fpg.unc.edu/resources/creating-culture-using-data-local-level

---

Data Workshop on Child Outcomes

Description: Based on a data workshop at the 2011 Measuring and Improving Child and Family Outcomes Conference, this is a series of narrated presentations with activities on using child outcomes data. The content and activities are:

- In first segment in the series, the progress category and summary statement data from a hypothetical state are compared with national data and data from states with a similar population size or percentage served.

Location: https://unc-fpg-cdi.adobeconnect.com/_a992899727/datawkshp1/

- The second segment digs a little deeper into the quality of state data by reviewing statewide missing data and trend data across years in order to answer the following questions: (1) Do we have enough data to trust the findings? and (2) Are the data stable?

Location: https://unc-fpg-cdi.adobeconnect.com/_a992899727/ecodatawkshop2/

Using Outcomes Data for Program Improvement

Description: This narrated PowerPoint presentation describes key concepts in using outcomes data for program improvement and highlights technical assistance materials developed by the Early Childhood Outcomes Center on how child outcomes data can be used at the state and local levels to improve programs. The session also reviewed the sections of the ECO state self-assessment that address using data for program improvement and showed how states and local programs can use the tool to chart their own progress toward data-based decision making.

Location: https://unc-fpg-cdi.adobeconnect.com/_a992899727/p3dzb60rtyq/

Conclusions

The Child Outcome Summary process was developed for reporting federally required data on outcomes experienced by children who receive early intervention or early childhood special education services. Accurately capturing the functional levels of children in these populations is challenging because they have a variety of delays and disabilities. The population of young children with disabilities contains far more developmental diversity than a typically developing population. Furthermore, no assessment tools have been developed to measure functioning on the three child outcomes states are required to report on. Reflecting recommended practice in early childhood for using multiple sources of information, the COS is a systematic process for summarizing multiple sources of information to produce a single rating for each of the three outcomes. The COS is not an assessment tool, nor was it designed to produce detailed information at the individual child level. It was designed to provide aggregated data for accountability and program improvement. The data produced by the COS process need to be sensitive enough to enable states to annually examine their performance with respect to a target and to look at the performance of local programs. The data also are used by the U.S. Department of Education to provide a national answer to the question of whether EI and ECSE are making a difference in children's lives.

Given the policy and programmatic significance of these uses of the data, there was a pressing need to establish the validity of the data produced by the COS process. The ENHANCE project carried out four separate studies to address this need. The collective findings from these studies indicate that the COS process produces data that can be effectively used for measuring child outcomes for accountability and program improvement. The studies also found that the states need to provide support for more consistent implementation and to strengthen their quality assurance activities to further improve the validity of the data. The findings from the studies were used to inform the revision of existing materials and the development of new materials to assist states in training providers in the COS process and in analyzing their COS data. Despite the implementation challenges uncovered, the body of evidence from this research indicates that the COS process is capable of producing data that are valid for the purposes for which it was developed.

References

American Academy of Pediatrics. (2001). Developmental surveillance and screening of infants and young children. *Pediatrics, 108*, 192-196.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Bagnato, S. J., Matesa, M., Smith-Jones, J., & Fevola, A. (2004). *The evidence for clinical judgment in early intervention*. Pittsburgh, PA: UCLID Center at the University of Pittsburgh. Retrieved from http://www.uclid.org:8080/uclid/pdfs/clinical_judgement.doc

Bagnato, S. J., Smith-Jones, J., Matesa, M., & McKeating-Esterle, E. (2006). Research foundations for using clinical judgment (informed opinion) for early intervention eligibility determination. *Cornerstones, 2*(3), 1-14.

Bailey, D. B., Buysse, V., Simeonsson, R. J., Smith, T., & Keyes, L. (1995). Individual and team consensus ratings of child functioning. *Developmental Medicine and Child Neurology, 37*(3), 246-259.

Bailey, D. B., Simeonsson, R. J., Buysse, V., & Smith, T. (1993). Reliability of an index of child characteristics. *Developmental Medicine and Child Neurology, 35*(9), 806-815.

Bailey, D. B. J., Bruder, M. B., Hebbeler, K., Carta, J., Defosset, M., Greenwood, C., . . . Barton, L. (2006). Recommended outcomes for families of young children with disabilities. *Journal of Early Intervention, 28*, 227-243.

Beckman, P. J. (1984). Perceptions of young children with handicaps: A comparison of mothers and program staff. *Mental Retardation, 22*, 176-181.

Bricker, D., & Squires, J. (1989). The effectiveness of parental screening of at-risk infants: The infant monitoring questionnaires. *Topics in Early Childhood Special Education, 9*(3), 67-85.

Bronfenbrenner, U., & Morris, P. (2006). *Handbook of Child Psychology* (R. M. Lerner Ed. 6 ed. Vol. 1). Hoboken, New Jersey: John Wiley & Sons, Inc.

Buysse, V., Smith, T., Bailey, D. B., & Simeonsson, R. J. (1993). Consumer validation of an index characterizing the functional abilities of young children with disabilities. *Journal of Early Intervention, 12*(3), 224-238.

Cairns, R. B., & Cairns, B. D. (2006). The making of developmental psychology. In R. M. Lerner (Ed.), *Handbook of child psychology* (6 ed., Vol. 1, pp. 89-165). Hoboken, NJ: John Wiley & Sons.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation in the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Council of Chief State School Officers. (2004). *A framework for examining validity in state accountability systems*. Washington, DC: Author.

Division for Early Childhood. (2007). *Promoting positive outcomes for children with disabilities: Recommendations for curriculum, assessment, and program evaluation*. Missoula, MT: Author.

Division for Early Childhood. (2014). *DEC recommended practices in early intervention/early childhood special education 2014*. Missoula, MT: Author. Retrieved from http://www.dec-sped.org/recommendedpractices

Early Childhood Outcomes Center. (2005). *Family and child outcomes for early intervention and early childhood special education*. Menlo Park, CA: Author. Retrieved from http://www.fpg.unc.edu/~eco/pdfs/eco_outcomes_4-13-05.pdf

Early Childhood Technical Assistance Center. (2014). *Draft system framework for Part C and Section 619*. Chapel Hill, NC: Author.

Fast, E. F., & Hebbler, S. (2004). *A framework for examining validity in state accountability systems. A paper in the series: Implementing the state accountability system requirements under the No Child Left Behind Act of 2001*. Washington, DC: Council of Chief State School Officers.

Gould, S. R., & Walker, D. (2008).*The process of Child Outcomes Summary Form ratings: Demonstrating progress.* Poster presented at the Division for Early Childhood 24th Annual International Conference on Young Children with Special Needs and Their Families.

Hebbeler, K., Barton, L., & Mallik, S. (2008). Assessment and accountability for programs serving young children with disabilities. *Exceptionality, 16*(1), 48−63.

Hebbeler, K., & Kahn, L. (2014). *What practitioners need to know about measuring EI and ECSE outcomes*. Paper presented at the Division for Early Childhood Conference, Philadelphia, PA.

Hebbeler, K., Spiker, D., Bailey, D., Scarborough, A., Mallik, S., Simeonsson, R., . . . Nelson, L. (2007). *Early intervention for infants and toddlers with disabilities and their families: Participants, services, and outcomes* (NEILS Final Report). Menlo Park, CA: SRI International.

Hebbeler, K. M., & Barton, L. R. (2007). The need for data on child and family outcomes at the federal and state levels. *Young Exceptional Children*, 1-15.

Hogan, C. (2001). *The power of outcomes: Strategic thinking to improve results for our children, families, and communities*. Washington, DC: National Governors Association.

Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin, 112*, 527-535.

Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Education measurement* (4th ed., pp. 17-64). Westport, CT: American Council on Education/Praeger Publishers.

Lazara, A., Danaher, J., Kraus, R., & Goode, S. (Eds.). (2009). *Section 619 profile* (16th ed.). Chapel Hill, NC: The University of North Carolina, FPG Child Development Institute, National Early Childhood Technical Assistance Center.

LeLaurin, K. (1990). Judgment-based assessment: Making the implicit explicit. *Topics in Early Childhood Special Education, 10*(3), 96-110.

McWilliam, R. A. (2004). DEC recommended practices: Interdisciplinary models. In S. Sandall, M. L. Hemmeter, B. J. Smith & M. E. McLean (Eds.), *DEC recommended practices: A comprehensive guide for practical application in early intervention/early childhood special education* (pp. 127-131). Longmont, CO: Sopris West.

Meisels, S. J. (2006). *Accountability in early childhood: No easy answers* (Occasional Paper 6). Chicago, IL: Herr Research Center for Children and Social Policy, Erikson Institute.

Meisels, S. J. (2007). Accountability in early childhood: No easy answers. In R. C. Pianta, M. J. Cox & K. L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 31-47). Baltimore, MD: Brookes Publishing.

Mislevy, R. J. (2006). Cognitive psychology and educational assessment. In R. L. Brennan (Ed.), *Education measurement* (4th ed., pp. 257-305). Westport, CT: American Council on Education/Praeger Publishers.

Morley, E., Vinson, E., & Hatry, H. P. (2000). *A look at outcome measurement in nonprofit agencies*. Washington, DC: Urban Institute.

Moss, P. A., Girard, B. J., & Haniford, L. C. (2006). Validity in education assessment. *Review of Educational Research, 30*(1), 109-162.

National Association for the Education of Young Children. (2009). *Developmentally approrpriate practice in early childhood programs serving children from birth through age 8* (Position Statement). Washington, DC: Author.

National Association for the Education of Young Children, & National Association Early Childhood Specialists in State Departments of Education. (2003). *Early childhood curriculum, assessment and program evaluation*. Washington, DC: Author.

National Association of School Psychologists (2005). *Position statement on early childhood assessment*. Bethesda, MD: Author. Retrieved from http://www.nasponline.org/information/pospaper_eca.html

National Early Childhood Accountability Task Force. (2007). *Taking stock: Assessing and improving early childhood learning and program quality*.

National Research Council. (2001). *Educating children with autism*. Washington, DC: National Academies Press.

National Research Council. (2008). *Early childhood assessment: Why, what, and how?* Committee on Developmental Outcomes and Assessments for Young Children, C.E. Snow & S.B. Van Hemel, (Eds.). Washington, DC: The National Academies Press.

National Research Council and Institute of Medicine. (2000a). *After-school programs to promote child and adolescent development: Summary of a workshop*. Washington, DC: National Academies Press.

National Research Council and Institute of Medicine. (2000b). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.

Neisworth, J. T., & Bagnato, S. J. (2005). DEC recommended practices: Assessment. In S. Sandall, M. L. Hernandez, B. J. Smith & M. E. McLean (Eds.), *DEC recommended practices: A comprehensive guide for practical application in early intervention/early childhood special education* (pp. 45–70). Longmont, CO: Sopris West.

Newborg, J. (2005). *Battelle developmental inventory, second edition examiner's manual*. Itasca, IL: Riverside.

Osbourne, D., & Gaebler, T. (1992). *Reinventing government: How the entrepreneurial spirit is transforming the public sector*. Reading, MA: Addison-Wesley.

President's Commission on Excellence in Special Education. (2002). *A new era: Revitalizing special education for children and their families*. Washington, DC: U.S. Department of Education.

Senate Committee on Governmental Affairs. (1993). *Government performance results act of 1993 report*. Washington, DC: Author. Retrieved from http://www.whitehouse.gov/omb/mgmt-grpa/grprptm

Shaw, J. G., Hammer, D., & Leland, H. (1991). Adaptive behavior of preschool children with developmental delays: Parent versus teacher ratings. *Mental Retardation, 29*, 49-53.

Siegel, D. J. (2012). *The developing mind: How relationships and the brain interact to shape who we are* (Second ed.). New York, NY: Guilford Press.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.

Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland adaptive behavior scales: Second edition (Vineland II), Survey interview form/Caregiver rating form*. Livonia, MN: Pearson Assessments.

Suen, H. K., Logan, C. R., Neisworth, J. T., & Bagnato, S. J. (1995). Parent-professional congruence: Is it necessary? *Journal of Early Intervention, 19*(3), 243-252.

Suen, H. K., Lu, C.-H., Neisworth, J. T., & Bagnato, S. J. (1993). Measurement of team decision-making through generalizability theory. *Journal of Psychoeducational Assessment, 11*(2), 120-132.

Tape, T. G. (n.d.). *Interpreting diagnostic tests*   Retrieved from http://gim.unmc.edu/dxtests/Default.htm

The Early Childhood Outcomes Center. (2006). *Introduction to the ECO "crosswalks" of birth-to-five assessment instruments to early childhood outcomes*: Author. Retrieved from http://www.fpg.unc.edu/~eco/assets/pdfs/Crosswalk_intro.pdf

Thompson, B. (Ed.). (2003). *Score reliability. Contemporary thinking on reliability issues*. Thousand Oaks, CA: Sage Publications.

Thompson, R. A. (2006). The development of the person: Social understanding, relationships, conscience, self. In N. Eisenberg (Ed.), *Handbook of child psychology* (6 ed., Vol. 3, pp. 24-98). Hoboken, NJ: John Wiley & Sons.

U.S. Office of Management and Budget. (2006). *ExpectMore.gov: IDEA special education grants for infants and families*. Washington, DC: Author. Retrieved from http://www.whitehouse.gov/omb/expectmore/summary.10000190.2005.html

Appendices

A   COS Materials

B   Selected Resources Related to Outcomes Constructs

C   State Requirement Materials from the U.S. Department of Education

D   ENHANCE Provider Survey

E   Provider Survey Data for EI and ECSE

F   Psychometric Characteristics of the BDI-2 and Vineland-II

G   Crosswalks of BDI-2 and Vineland-II to Child Outcomes

H   ENHANCE Child and Family Information Form

I   Additional Descriptive Information about Longitudinal Sample

J   Child Assessments Study: Regressions Predicting COS Ratings

K   Levels of 7-Point COS Ratings with Mean Scores on BDI-2, Vineland-II, and ABILITIES Index at Program Entry in the Child Assessments Study

L   Team Decision-Making Study Project Developed Forms

M   Team Decision-Making Study: Regressions Predicting COS Ratings

N   ENHANCE Project Dissemination

Appendix A

COS Materials

Overview of the Child Outcomes Summary Process

Instructions for Completing the Child Outcomes Summary Form

Child Outcomes Summary Form

Decision Tree for Summary Rating Discussions

Definitions for Outcome Ratings

Documentation Key for Outcome Ratings

At a Glance: Child Outcomes

Overview of the
Child Outcomes Summary Process

The following questions and answers reflect the ECO Center's current thinking on the use of the **Child Outcomes Summary (COS) Process**. We continue to receive feedback on the COS process and the supporting materials and have made revisions based on that feedback. Please continue to check the ECO web site for additional revisions.

1. What is the COS process?

The COS process is a team process for summarizing information related to a child's progress on each of the three child outcome areas on a 7-point scale. The COS process can be used:

1) When the state wants to use multiple sources of information to describe a child's functioning on each of the outcomes. The information could include one or more norm-referenced or curriculum-based assessments, parent report on child's skills and behavior, progress notes of therapists working with the child, observations by a teacher or child care provider, or other sources; and/or

2) When different assessments have been given to different children across the state and the results need to be placed on the same scale to be aggregated.

The COS form is **NOT** an assessment instrument. It is a document used for summarizing across multiple sources of information about the child. The COS process allows states to address the OSEP reporting requirement as well as look at the child outcomes data in other ways. Using the COS process does not require that programs collect more data about children's progress; it is a mechanism that allows them to summarize assessment information for federal reporting as well as for their own purposes, such as for accountability, program planning, and program improvement.

2. What materials related to the COS process are available on the ECO web site?

The following materials are available on the ECO web site:

- *Instructions for Completing the Child Outcomes Summary (COS) form* – This document contains instructions for completing the form along with the definitions of the scale points.
- *Child Outcomes Summary (COS) form* **–**. This form can be used to summarize information from multiple sources on the 3 outcomes. The form includes a cover sheet and 3 pages, one for each outcome.
- *How Data from the COS form Can be Used to Address the OSEP Reporting Requirement* – This document explains how data from the COS form at entry and at exit produces data required by OSEP.
- *COS to OSEP Categories Calculator- Model 2.0* **–** This excel file demonstrates how various combinations of outcome ratings at entry and exit are converted into each of the 5 OSEP reporting categories. The new version (Model 2.0) also automatically creates graphic presentations (charts and tables) of entry and exit data and progress

categories.

- **_Child Outcome Summary (COS) Form Training Materials_** – Materials include a sample training agenda and power point presentations and activities on a number of COS topics (e.g. COS refresher, working with families, COS quality assurance, looking at data). Materials can be adapted for state training sessions.
- **_State-Developed COS Materials_**- These COS materials were created or adapted by a number of states. These materials have not been reviewed or critiqued by the ECO center, but are provided as a convenience for others to use to meet training needs in their specific locations.

3. What is the basis for the scale on the COS form?

The scale on the COS form is based on several assumptions:

> a. The overall goal of programs and services for children is active and successful participation now and in the future across a variety of settings. Achieving each of the three outcomes is key to this overall goal.
> b. For many, but certainly not all young children with disabilities, receipt of high quality services will allow them to move closer to age-appropriate functioning than they would have been able to without those services.
> c. Documenting children's movement toward age-appropriate functioning is one type of evidence that can be used to make a case for the effectiveness of early intervention and early childhood special education.

Building off of these assumptions, the highest end of the scale represents age-expected or age-appropriate functioning with each lower point being a degree of distance from age expectations. Additional information about the scale points is included in the document *"Instructions for Completing the Child Outcome Summary (COS) Form."*

4. Who completes the COS process?

States need to decide who completes the COS process. The ECO Center recommends that the ratings be determined by a team including family members, professionals who work with the child, and others familiar with the child's functioning. Teams in states already using the COS process generally range from 2 -7 people. The ECO Center strongly recommends that the ratings not be determined by individuals who do not know the child, for example, by assigning a rating based only on information available in the child's records, or for the rating scale to be provided without any guidance or instructions.

5. How often is the COS process to be completed?

To provide data for the OSEP reporting requirements, the COS process must be completed at a minimum once at program entry and again at program exit with at least 6 months in between. States that want outcome data for their own purposes should consider completing the form more often, for example, annually or every 6 months.

6. When is the COS process to be completed?

The COS process is to be completed in present time to reflect the child's current functioning. Some states are completing the form at IFSP or IEP meetings and regularly scheduled reviews

when a team is assembled but it does not have to be done as part of one of these meetings. We recommend against trying to assign ratings for past time periods, for example, trying in December to assign a rating for what the child was like in August.

7. *Are training materials available*?

The ECO Center has developed training and guidance materials for use of the COS process. Materials are available on the ECO website under Professional Development Resources. These include a sample training agenda, power point presentations and activities related to both general outcomes topics as well as COS topics. We always welcome suggestions for training materials as well as information about training approaches underway in states and programs.

8. *How much training is required to use the COS process?*

Field testing has shown that training is essential to effective and reliable use of the COS process. Sample training materials available on the ECO website outline a day-and-a-half training activity that includes general background information on child outcomes measurement, information to promote understanding of the three outcome areas, and references to recommended assessment practices. In addition, field testing has shown that training must include opportunities to "practice" the ratings through case examples. Walking through several cases in a large group discussion and smaller team breakout formats help to clarify differences between the points on the rating scale, as participants review multiple sources of assessment information about a child, compare that information to age expectations, and then determine a rating. We also strongly recommend that states develop opportunities for periodic feedback sessions with providers, after they have begun to use the summary form. These sessions will allow individuals to share effective strategies that work with specific local populations as well as to ask questions and share information in an effort to enhance the consistency of approaches used.

9. *Is information available on the validity and reliability of the COS ratings?*

The ECO Center is currently funded to examine the validity of the COS ratings. Ideally, this information would have been available before the tool was released but the OSEP reporting timeline did not allow for this. Preliminary results from pilot data collected to date is promising, leading us to feel confident that under proper conditions (e.g., sufficient training, adequate opportunities to have questions addressed, adequate monitoring of the process), the COS will produce valid and reliable information. Data and future findings will be shared on the ECO website as they become available.

10. *Why is the COS copyrighted?*

States and programs are encouraged to use and reproduce the form. There is no charge to use any materials produced by the ECO Center. We are copyrighting materials to prevent anyone from charging for them in the future.

11. *Why should we contact ECO if we want to use or adapt the COS process?*

ECO would like to keep track of which states are using the form to learn more about how the process is working. We would appreciate a state contacting us so we have an accurate list. At

that time we will also be happy to discuss advantages and disadvantages of potential adaptations to the COS process that states are considering.

12. Can a state make changes to the COS process or form?

States can change the process or form to meet their needs but we encourage them to think through the consequences of those changes. Some adaptations, such as formatting changes, are minor and not likely to impact the type of data that will result. Other adaptations (for example, using a 5-point instead of a 7-point scale) are major and could interfere with easy translation from the rating scale to the OSEP reporting categories. Using a very different scale also will mean that the lessons being learned from piloting the ECO version of the form will not apply. We plan to collect considerable data about the use and properties of the summary process and it will be difficult to say how much of that research will apply if a state has made major changes in the COS process or form. We encourage states to contact us to discuss what they would like to change so we can keep track of which states are using the ECO-developed process and form and which are using their own adaptation.

13. What if I have a question about the COS process or the related materials? Can states and others comment on the COS process and related materials?

Yes. Questions and comments are encouraged. Send them to staff@the-eco-center.org .

Instructions for Completing the Child Outcomes Summary Form

Directions for Completing the Form

1. Page 1: Provide all the requested information. It is strongly recommended that the family be asked to provide information about the child's functioning, but if the family's information was not included, check "not included." Additional state-specific information also may be requested.

2. Questions 1a, 2a, 3a: Circle only **one** number for each outcome. Definitions for the scale points are provided at the end of the instructions.

3. Supporting evidence: Provide the evidence that supports the rating. Indicate the source of the evidence (e.g., parent, speech therapist, teacher, XYZ assessment) and the nature of the evidence from the source. For example, if a child's functioning receives a rating of '5', relevant results should provide evidence of a mix of age appropriate and not age appropriate skills and behaviors. A sample completed evidence table is provided below.

Source of information	Date	Summary of Relevant Results
Candace's mom	4/12/06	Mom reports that when Candace eats by herself she makes a big mess. She eats finger foods but does not use a fork or spoon. She uses a "sippy" cup with two hands. Mom reports that she has not begun to toilet train Candace. Candace does not let mom know when she has a wet or soiled diaper. She pulls off her socks when getting ready for bed.
Candace's child care provider	4/5/06	Child care provider said that Candace is learning to use a spoon, but usually uses her fingers to feed herself. Candace uses diapers and tugs on diaper after it is wet or soiled.
Carolina Curriculum for Infants and Toddlers with Special Needs	Administered 3/13/06	Self-Help: Eating – 12-15 months Self-Help: Dressing – 15-18 months Self-Help: Grooming – 18-21 months Self-Help: Toileting -- <15-18 months
Developmental specialist	Observed over a 4 week period in March 2006	Observed in her child care environment during structured activities and unstructured play time. She clapped and jumped during a group song. During free play Candace tended to sit quietly unless engaged in a play activity by her caregiver. Candace did not object to having hands washed by caregiver, but needed assistance.

4. Questions 1b, 2b, 3b: Complete questions 1b, 2b, and 3b only when questions 1a, 2a, and 3b have been answered previously. Circle one number to indicate if the child has made progress since the previous outcomes rating. Progress is defined as the acquisition of at least one new skill or behavior related to the outcome. Describe the general nature of the progress in the space provided.

To Help You Decide on the Summary Rating for Questions 1a, 2a, and 3a:

This outcomes summary asks you to consider and report on what is known about how this child behaves across a variety of settings and situations.[*] Children are with different people (for example, mother, big brother, child care provider) and in different settings (for example, home, grocery store, playground). The summary rating provides an overall picture of how the child behaves across the variety of people and settings in his or her life at this particular time in his or her life.

In addition to summarizing across settings and situations, the rating process asks you to compare a child's skills and behaviors to those of his or her same-age peers. For each of the three summary questions, you need to decide the **extent to which the child displays behaviors and skills expected for his or her age** related to each outcome area.

The summary scale is based on a developmental framework that assumes:

1. Children develop new skills and behaviors and integrate those skills and behaviors into more complex behaviors as they get older;
2. These skills and behaviors emerge in a somewhat predictable developmental sequence in most children, thus allowing for descriptions of what 2 year olds generally do, what 3 year olds generally do, etc.;
3. The development of children with disabilities can be compared to the development of their same-age peers.
4. Some of the skills and behaviors that develop early serve as the foundation for later skills and behavior, or expressed another way, later skills build on earlier skills in predictable ways. Teachers and therapists can use the earlier skills to help children move to the next higher level of functioning developmentally. We refer to these earlier skills that serve as the base and are conceptually linked to the later skills, as "**immediate foundational skills**." For example, children play along side one another before they interact in play.
5. Some children's development is characterized by delays, meaning they acquire skills and behaviors at a substantially slower pace than other children.
6. Some children's development is atypical in that their functioning is so different from that of other children their age that it is considered outside the limits of age expected behavior for children of that age.

Use the following information to help you answer each question:

- Ratings are expected to take into account the child's functioning across a full range of situations and settings. Therefore, information from many individuals in contact with the child could be considered in deciding on a rating. These may include (but are not limited to): parents and family members, caregivers or child care providers, therapists, service providers, case managers, teachers, and physicians. If there is not enough information available about a child's functioning across settings and situations, you will need to gather more information before you can decide on a rating.

- Many types of information could be considered in selecting a rating. These may include (but are not limited to): parent and clinical observation, curriculum-based

---

[*] Note: The outcomes summary form was not designed to determine eligibility for services. It would be inappropriate to use it in this way.

assessments, norm-referenced assessments, service provider notes about performance in different situations, and progress and issues identified in the IFSP/IEP or individualized planning process.

- Depending on the assessment tool, assessment tools can be a useful source of information for reaching a summary decision but resulting information should be placed in context with other information available about a child. Many assessment tools are domain-based and were not designed to provide information about functional behaviors and functioning across a variety of situations. Knowing that a child has or has not mastered assessment items that are related to the outcome provides helpful information but the information should be used in conjunction with what else is known about the child. A high score on a set of items in a domain related to the outcome might not mean the child has achieved the outcome and, conversely, a low score might not mean the child has not achieved it.

- Ratings should reflect the child's current functioning across settings and in situations that make up his/her day. Ratings should convey the child's functioning across multiple settings and in everyday situations, *not* his/her capacity to function under unusual or ideal circumstances.

- A standardized testing situation is an unusual setting for a young child. If the child's functioning in a testing situation differs from the child's everyday functioning, the rating should reflect the child's everyday functioning.

- If the child is from a culture that has expectations that differ from published developmental milestones for when young children accomplish common developmental tasks, such as feeding themselves or dressing themselves, use the expectations for the child's culture to decide if child's functioning is at the level expected for his or her age.

- If the child was born prematurely, use the expectations for the child's chronological age, not the corrected age. The intent of the form is to describe the child's current functioning relevant to expectations for his or her age. Presumably over time and with support, many children born prematurely eventually will perform like same age peers.

- If assistive technology or special accommodations are available in the child's everyday environments, then the rating should describe the child's functioning using those adaptations. However, if technology is only available in some environments or is not available for the child, rate the child's functioning with whatever assistance is commonly present. Ratings are to reflect the child's **actual** functioning across a range of settings, *not* his/her capacity to function under ideal circumstances if he or she had the technology.

Additional Information

The outcomes reflect several beliefs about young children:

- It is important that all children be successful participants in a variety of settings both now and in the future. Achieving the three outcomes is key to being successful participants in life.

- Programs for young children and their families are working to ensure that all children will have the best possible chance of succeeding in kindergarten and later in school – even though school might be several years off for some children. Children who have achieved the outcomes at a level comparable to their same aged peers prior to kindergarten entry have a high probability of being successful in kindergarten.

- Learning and development occur continuously in the years preceding kindergarten. There is much variation in how children develop but children whose development is consistently below what is expected for their age are at risk of not being successful in kindergarten and later school years.

CHILD OUTCOMES SUMMARY FORM

Date: _____/_____/_____
            Mon   Day   Yr

Child Information

Name: _____

Date of birth: _____/_____/_____
               Mon   Day   Yr

ID:_____

Persons involved in deciding the summary ratings:

Name	Role

Family information on child functioning (Check all that apply):

___ Received in team meeting

___ Collected separately

___ Incorporated into assessment(s)

___ Not included

1. POSITIVE SOCIAL-EMOTIONAL SKILLS (INCLUDING SOCIAL RELATIONSHIPS)

To answer the questions below, think about the child's functioning in these and closely related areas (as indicated by assessments and based on observations from individuals in close contact with the child):

- *Relating with adults*
- *Relating with other children*
- *Following rules related to groups or interacting with others (if older than 18 months)*

1a. To what extent does this child show age-appropriate functioning, across a variety of settings and situations, on this outcome? *(Circle one number)*

Not Yet		Emerging		Somewhat		Completely
1	2	3	4	5	6	7

Supporting evidence for answer to Question 1a

Source of information	Date	Summary of Relevant Results

1b. *(If Question 1a has been answered previously):* **Has the child shown *any* new skills or behaviors related to positive social-emotional skills (including positive social relationships) since the last outcomes summary?** *(Circle one number)*

Yes	1→	Describe progress:
No	2	

2

2. ACQUIRING AND USING KNOWLEDGE AND SKILLS

To answer the questions below, think about the child's functioning in these and closely related areas (as indicated by assessments and based on observations from individuals in close contact with the child):

- *Thinking, reasoning, remembering, and problem solving*
- *Understanding symbols*
- *Understanding the physical and social worlds*

2a. To what extent does this child show age-appropriate functioning, across a variety of settings and situations, on this outcome?   *(Circle one number)*

Not Yet		Emerging		Somewhat		Completely
1	2	3	4	5	6	7

Supporting evidence for answer to Question 2a

Source of information	Date	Summary of Relevant Results

2b. *(If Question 2a has been answered previously):* **Has the child shown *any* new skills or behaviors related to acquiring and using knowledge and skills since the last outcomes summary?**   *(Circle one number)*

Yes	1→	Describe progress:
No	2	

3

3. TAKING APPROPRIATE ACTION TO MEET NEEDS

To answer the questions below, think about the child's functioning in these and closely related areas (as indicated by assessments and based on observations from individuals in close contact with the child):

- *Taking care of basic needs (e.g., showing hunger, dressing, feeding, toileting, etc.)*
- *Contributing to own health and safety (e.g., follows rules, assists with hand washing, avoids inedible objects) (if older than 24 months)*
- *Getting from place to place (mobility) and using tools (e.g., forks, strings attached to objects)*

3a. To what extent does this child show age-appropriate functioning, across a variety of settings and situations, on this outcome?   *(Circle one number)*

Not Yet		Emerging		Somewhat		Completely
1	2	3	4	5	6	7

Supporting evidence for answer to Question 3a

Source of information	Date	Summary of Relevant Results

3b. *(If Question 3a has been answered previously):* **Has the child shown *any* new skills or behaviors related to taking appropriate action to meet needs since the last outcomes summary?**   *(Circle one number)*

Yes	1→	Describe progress:
No	2	

4

Decision Tree for Summary Rating Discussions

Does the child ever function in ways that would be considered age-appropriate with regard to this outcome?

No (consider rating 1-3)

Yes (consider rating 4-7)

Does the child use any immediate foundational skills related to this outcome upon which to build age-appropriate functioning across settings and situations?

Is the child's functioning age-appropriate across all or almost all settings and situations?

No

Yes

No

Yes

To what extent is the child using immediate foundational skills across settings and situations?

To what extent is the child's functioning age-appropriate across settings and situations?

Does anyone have concerns about the child's functioning with regard to the outcome area?

Uses skills that are not yet immediate foundational

Occasional use of immediate foundational skills

Uses immediate foundational skills most or all of the time across settings and situations

Occasional use of age-appropriate skills; more behavior that is not age-appropriate

Uses a mix of age-appropriate and not age-appropriate behaviors and skills across settings and situations

Yes

No

Rating = 1

Rating = 2

Rating = 3

Rating = 4

Rating = 5

Rating = 6

Rating = 7

Definitions for Outcome Ratings:

For Use with the Child Outcomes Summary (COS) Form

Overall Age-Appropriate	**Completely** *means*:	**7**	• Child shows functioning expected for his or her age in **all or almost all everyday situations** that are part of the child's life. Functioning is considered **appropriate** for his or her age. • No one has any concerns about the child's functioning in this outcome area.
		6	• Child's functioning generally is considered **appropriate** for his or her age but there are **some significant concerns** about the child's functioning in this outcome area. These concerns are substantial enough to suggest monitoring or possible additional support. • Although age-appropriate, the child's functioning may border on not keeping pace with age expectations.
Overall Not Age-Appropriate	Somewhat *means:*	**5**	• Child shows functioning expected for his or her age **some of the time and/or in some settings and situations**. Child's functioning is a mix of age-appropriate and not age-appropriate behaviors and skills. • Child's functioning might be described as like that of a **slightly younger child***.
		4	• Child shows occasional age-appropriate functioning across settings and situations. More functioning is **not** age-appropriate than age-appropriate.
	Nearly *means:*	**3**	• Child does **not yet** show functioning expected of a child of his or her age in any situation. • Child uses **immediate foundational skills,** most or all of the time, across settings and situations. Immediate foundational skills are the skills upon which to build age-appropriate functioning. • Functioning might be described as like that of a **younger child***.
		2	• Child occasionally uses **immediate foundational skills** across settings and situations. More functioning reflects skills that are **not** immediate foundational than are immediate foundational.
	Not yet *means*:	**1**	• Child does **not yet** show functioning expected of a child his or her age in any situation. • Child's functioning does **not yet include immediate foundational skills** upon which to build age-appropriate functioning. • Child functioning reflects skills that developmentally come before immediate foundational skills. • Child's functioning might be described as like that of a **much younger child***.

* The characterization of functioning like a younger child only will apply to some children receiving special services, such as children with developmental delays.

The Early Childhood Outcomes Center 11/8/12

Documentation Key for Outcome Ratings:
For Use with the Child Outcomes Summary Form (COSF)*

<table>
<tr>
<td rowspan="2">Overall Age-Appropriate</td>
<td>Completely</td>
<td>7</td>
<td>
• Provide examples of child's age-appropriate functioning

• Indicate: "no concerns"
</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>
• Provide examples of the child's age-appropriate functioning.

• Note concerns

• <i>If there is evidence of functioning that is not age appropriate, a rating of 6 or 7 should not be assigned</i>
</td>
</tr>
<tr>
<td rowspan="5">Overall Not Age-Appropriate</td>
<td>Somewhat</td>
<td>5</td>
<td>
• Provide examples of child's age-appropriate functioning

• Provide examples of the child's functioning that is NOT age appropriate
</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>
• Provide examples of age appropriate functioning

• Provide examples of the child's functioning that is NOT age-appropriate

• <i>Evidence should show more functioning that is NOT age appropriate</i>
</td>
</tr>
<tr>
<td>Nearly</td>
<td>3</td>
<td>
• Provide examples of child's functioning at the immediate foundational skill level **

• <i>No age appropriate functioning should be noted for a rating of 3</i>
</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>
• Provide a few examples of the child's functioning at the immediate foundational skill level

• Provide examples of the child's functioning that is not yet age appropriate or immediate foundational

• <i>Evidence should show more functioning that is NOT immediate foundational than is immediate foundational for a rating of 2</i>
</td>
</tr>
<tr>
<td>Not yet</td>
<td>1</td>
<td>
• Provide examples of the child's functioning that is not yet age appropriate or immediate foundational skills

• <i>No age-appropriate or immediate foundational functioning should be noted for a rating of a 1</i>
</td>
</tr>
</table>

** Immediate foundational skills are the skills upon which to build age-appropriate functioning.

*Adapted by Mecklenburg County Children's Developmental Services, Charlotte, NC, 6/8/09

The Early Childhood Outcomes Center   **5/21/09**

At a Glance: Child Outcomes

Child Outcomes

States are required to report on the percentage of infants and toddlers with Individualized Family Service Plans (IFSPs) or preschool children with Individualized Education Plans (IEPs) who demonstrate improved:

1. Positive social-emotional skills (including social relationships);
2. Acquisition and use of knowledge and skills (including early language/communication [and early literacy]); and
3. Use of appropriate behaviors to meet their needs.

Positive social emotional skills (including social relationships). This outcome involves relating to adults, relating to other children, and for older children, following rules related to groups or interacting with others. The outcome includes concepts and behaviors such as attachment/separation/autonomy, expressing emotions and feelings, learning rules and expectations in social situations, and social interactions and social play.

Acquisition and use of knowledge and skills (including early language/communication/early literacy). This outcome involves activities such as thinking, reasoning, remembering, problem solving, number concepts, counting, and understanding the physical and social worlds. It also includes a variety of skills related to language and literacy including vocabulary, phonemic awareness, and letter recognition.

Use of appropriate behaviors to meet their needs. This outcome involves behaviors like taking care of basic needs, getting from place to place, using tools (such as forks, toothbrushes, and crayons), and, in children 24 months or older, contributing to their own health, safety, and well-being. It also includes integrating motor skills to complete tasks; taking care of one's self in areas like dressing, feeding, grooming, and toileting; and acting on the world in socially appropriate ways to get what one wants.

Ultimate goals for early intervention and early childhood special education:

For children…

to enable young children to be active and successful participants during the early childhood years and in the future in a variety of settings – in their homes with their families, in child care, preschool or school programs, and in the community.

For families …

to enable families to provide care for their child and have the resources they need to participate in their own desired family and community activities.

<u>**Child Outcomes are:**</u>
- **A snapshot –** of the child's overall functioning at one given point in time, informed by the full team of people who know the child best across settings and situations where the child spends his/her time. They provide a consistent format for programs to see the extent to which their activities are making a difference in supporting all children's progress and offer needed information to guide program improvement. More detailed information about each specific child's functioning and progress may be more useful to the team in developing program plans than the 3 global outcomes.
- **Integrated –** going beyond skills in any one domain, bringing them together in complex and interconnected ways
- **Functional –** reflecting how children use the skills they have in everyday activities to accomplish things that are meaningful to them. They go beyond actions that might be observed in a child sporadically under a specific set of ideal or unusual conditions to focus on how the child regularly uses his/her skills
- **Different across contexts –** in many cases children's functioning will vary across contexts, exhibiting different ways of interacting with different people and in places where different supports and expectations exist. The outcomes reflect an overall sense of how the child functions across the full range of everyday settings and situations. Include a child's functioning with whatever assistive technology supports may routinely be available (or not) in the settings where the child spends his/her time
- **Considered relative to same-age peers –** ratings reflect the child's functioning relative to that of same-aged peers to help interpret the mix of functioning observed and the trajectory of the child's progress over time
- **Not intended for eligibility determination –** the outcomes reflect one measure of a child's functioning. For a variety of reasons, it is not expected to mirror eligibility determination. A number of kids eligible for IDEA-funded services may demonstrate age-expected functioning in one or more of the outcome areas

<u>**Progress Categories**</u>
For OSEP, states are required to report on the percentage of children in five categories of progress for each of the three child outcomes (percentage in 5 categories X 3 outcomes = 15 numbers the state reports):

a. Children who did not improve functioning.
b. Children who improved functioning but not sufficient to move nearer to functioning comparable to same aged peers.
c. Children who improved functioning to a level nearer to same aged peers but did not reach it.
d. Children who improved functioning to reach a level comparable to same aged peers.
e. Children who maintained functioning at a level comparable to same aged peers.

States must report progress category information on children who receive services in the state for 6 months or more. Progress category information on all 3 outcomes is required for each child, regardless of the child's reason for eligibility. Categories are derived by combining the outcomes ratings or descriptor statements given at program entry and exit.

<u>**Summary Statements**</u>
For OSEP, states are required to convert information from the progress categories into two summary statement percentages for each of the three child outcomes:

<u>Summary Statement 1</u>: Of those children who entered the program below age expectations, the percent who substantially increased their rate of growth by the time they exited the program. (State derives a percentage for each child outcome area.) *Formula*: [(c + d)/(a + b + c + d)] x 100, where letters represent the actual number of children in each progress category group.

<u>Summary Statement 2</u>: The percent of children who were functioning within age expectations in each Outcome by the time they exited the program. (State derives a percentage for each child outcome area.) *Formula*: [(d + e)/(a + b + c + d + e)] x 100, where letters represent the actual number of children in each progress category group.

Appendix B

Selected Resources Related to Outcomes Constructs

Crosswalking Rules of Thumb

Discussion Prompts for Teams

Background:

    As part of technical assistance to state Part C and state Part B Preschool programs, the Early Childhood Outcomes (ECO) Center categorized the content on a number of commonly administered assessment tools into to the three child outcome areas. A full list of these crosswalks is available at: http://ectacenter.org/eco/pages/crosswalks.asp. The assessment tools used in ENHANCE's child assessments study can be found in Crosswalks of BDI-2 and Vineland-II, another appendix to this report.

    States use these crosswalks to help explain to providers what assessment content to reference during COS discussions about each outcome. The crosswalks also are a resource about how to use assessment data in team discussions.

    This appendix includes two specific resources. The first is the documentation used by the ECO Center describing rules of thumb for which skills and behaviors should be considered in discussion of each of the outcome areas. Specific details about how skills in each of the developmental domains relates to outcomes is considered. The second resource is a list of probing discussion questions to help teams think about the kinds of information they need to have available as they make COS decisions about each outcome area. By reviewing the list of discussion questions, teams get a sense for how topics relate to each outcome area. Teams are not expected to raise or discuss all of these questions. Rather the discussion prompts are a tool for professional development of those involved in the COS process.

Crosswalking Rules of Thumb

General Rules:

1. Level at which assessment tools are crosswalked

We crosswalked tools at the smallest level the instrument can be used and the smallest level that communicates content most effectively, i.e., sub-areas or items. For most criterion-referenced or curriculum-based assessment tools, we used the sub-area level with the developer's headings and often provided examples of items from that sub-area to illustrate the aspects of development that relate to the outcome.

Norm-referenced tests[1] always are crosswalked at the lowest level that the tool developers recommend for valid interpretation of the data and have provided normative information. This is usually at a subscale or sub-domain level.

In some crosswalks, headings appear under more than one child outcome. This occurs when skills and behaviors within a tool's sub-area address more than one outcome. For example, under the sub-area of language development, skills and behaviors related to conversational skills are listed under Outcome 1, skills and behaviors related to using language to get needs met are listed under Outcome 3, and skills and behaviors related to developing vocabulary and grammar are listed under Outcome 2. For more information about decision-making regarding double classification, see Section 4.

2. Assignment of assessment area or sub-area to an outcome

We placed skills and behaviors addressed in the assessment tool under the outcome to which they are most closely linked conceptually, e.g., skills and behaviors related to getting along with peers go with Outcome 1. We looked at each targeted skill and asked whether knowing how the child did in that skill area would help decide whether or not he or she had achieved the outcome. If a child is not yet proficient in a specific skill area, for example, does that provide important information about the child's outcome? Classification decisions were based on the description of the content that is the focus of the area rather than the heading title or specific information from examples in the manual.

We made an effort to understand the tool developer's intent when assigning assessment categories to outcomes. In other words, we considered the focus of the majority of skills and behaviors under a sub-area in order to judge whether the sub-area was most intended to address positive social relationships, acquisition and use of knowledge and skills, or taking action to meet needs. Since much of development is interrelated, many skill areas overlap and can be logically assigned to more than one outcome. For information guiding decisions about double classification, see section 4.

With the curriculum-based measures, we tried to avoid splitting sub-areas across outcomes. This could not be avoided, however, when some of the skills and behaviors included in an area

---

[1] Crosswalks of norm-referenced instruments include a note providing information about the lowest appropriate threshold for crosswalking on that specific instrument.

clearly addressed one outcome and others clearly addressed a different outcome. We never split sub-areas in the norm-referenced instruments.

Particularly in the sub-areas of language and learning, it is difficult to assign skills and behaviors to outcomes when item content lacks specificity. In such cases we elected to assume that the item pertained to a general, overarching acquisition and use of knowledge and skills, and therefore made the assignment to Outcome 2. For example, if a skill or behavior related to learning or communicating is presented without child's intent or purpose being addressed, it is placed with Outcome 2 (e.g., "asks questions" goes with Outcome 2, whereas "asks questions to get needs met" goes with Outcome 3).

We considered some skills to be forerunners or "precursors," in that their presence might allow further development of certain skills and behaviors. Precursor skills that are clearly linked to one of the outcomes were placed with the outcome. We assigned precursor skills that are general or cross-cutting and apply equally to all 3 outcomes to Outcome 2, as part of general acquisition and use of knowledge and skills.

3. Level of detail in the crosswalks

In general, we summarized targeted skills and behaviors at the sub-area level. Rather than include lots of items verbatim, in many tools, we included examples of items under each heading of the assessment tool to illustrate the aspects of development addressed by that sub-area. When example items are included, we list them in order from early development to later development.

When it was necessary to split a sub-area across outcomes, we provided more detail to show what skills and behaviors within the sub-area relate to each outcome.

4. Double classification

Targeted skills and behaviors were double classified when we felt that they contributed equally or nearly equally toward understanding achievement of more than one outcome.

We tried to minimize double classification because it creates lots of redundancy and interferes with differentiating the core skills that contribute to achieving each of the outcomes. Skills or behaviors that relate to a second outcome area, but not as strongly as they relate to a primary outcome area, were only classified with the prime area. Realistically, many targeted skills and behaviors could have been double or triple classified because of the interrelated nature of development in young children and the need to integrate skills across domains to achieve the functional outcome. However, we examined the focal contribution of each area to understanding the outcomes and classified the conceptual relationships accordingly.

5. Universal design and items that apply to some children, but not all children

There are many possible ways of reaching competence on the three functional outcomes. Assessment tools often tap skills and behaviors that are the most common developmental approaches (especially for fine and gross motor skills). However, certain items on a tool might not apply to all children with disabilities. Some children may never develop the skills and If items provide little information for children who will meet their needs in other ways, then they should not be considered in assessing the child's achievement of the outcomes. For example, if a child

moves from place to place to meet her needs through effective use of a wheel chair rather than using gross motor skills, her lack of independent walking does not provide useful information about whether or not she can take action to meet her needs. Items that do not reflect universal design, i.e., assume typical body functioning (speaking, seeing, movement, hearing) are marked with an asterisk. If a child has typical body functioning, these items provide useful information, but for children who do not, the lack of the skill is not necessarily an indicator of the level of attainment of the outcome. The note related to the asterisk reads: "Precursor skills for functional behaviors. These skills may not be appropriate or expected for some children, including those with sensory, motor, or other impairments."

Likewise, some tools include skills and behaviors that are specific to children with certain kinds of disabilities (e.g., section on wheelchair skills or sign language skills). We include these items in the crosswalk with a note, but they are only informative about the child's overall functioning on the outcome in cases where the child uses those skills as a major means to accomplish the given outcome. In other cases, performance on these skills contributes little to understanding the child's functioning on these the relevant outcome. The note listed on the crosswalk in these cases is: "Some items relate to assessment of specific skills that are most relevant for children with certain types of sensory, motor, or other impairments. These skills may not be appropriate or expected to contribute information about functioning on the outcome for many children."

6. Inclusion of every item in a crosswalk

Whenever possible we tried to link to outcomes all of the skills or behaviors included on an assessment tool. However, if proficiency or lack of proficiency on an item or sub-area of the tool did not provide information about the child's achievement of any of the three outcomes, then it was not classified. See the examples below.

The decision not to classify every item or area of an assessment does *not* mean that they are not important experiences for young children. It is only a statement that achieving or not achieving one of these items does not provide information about attainment of any of these three outcomes.

Areas of an assessment tool were left out of crosswalks when the majority of skills and behaviors included in that area did not contribute to understanding the child's functional abilities in any outcome area. However, if the area of an assessment included only a few non-relevant skills and behaviors, then we included that area on the crosswalk.

When skills or areas were not classified, they were listed in a footnote on the crosswalk. The footnote reads: "Note: Areas that are not precursor to or components of any of the three outcomes, and therefore not included in the crosswalk.'

On norm-referenced tests, a note is included if a subscale is crosswalked to an outcome area, but major portions of the subscale content are not relevant to the specific outcome area or would have been left out of the crosswalk altogether had a more refined level of specificity been possible. Footnote terminology is: "** This composite/subscale/area includes significant content that is not precursor to or components of the specified outcome."

Items/skills that have been left off of other crosswalks include the following.[2]
- Moving to music (High/Scope 0-3, 3-6),
- Communicating through rhythm (HELP); Feeling and expressing a steady beat (High/Scope 3-6)
- Moving in various ways (High/Scope 3-6)
- Moving with objects (High/Scope 0-3, 3-6)
- Singing (High/Scope 3-6)
- The arts: Expression and representation (Work Sampling)
- The arts: Understanding and appreciation (Work Sampling)
- Jumping (HELP 0-3, 3-6; Brigance IED-II)
- Hopping (Brigance IED-II)
- Catching (HELP 3-6; Brigance IED-II)
- Catching/Throwing (HELP 0-3)
- Catching/Trapping under Early Movement Indicator (IGDI)
- Throwing, kicking, and catching skills (Creative Curriculum)
- Throwing/Rolling (IGDI; Brigance IED-II)
- Bilateral play (HELP)
- Balance beam (HELP 0-3, 3-6; Brigance IED-II)
- Swimming (HELP)
- Blocks/puzzles (HELP 3-6); Block construction (HELP 0-3)
- Builds tower with blocks (Brigance IED-II)
- Formboard (HELP 0-3)
- Pegboard (HELP 0-3)
- Paper activities (HELP 0-3, 3-6)
- Stringing beads (HELP 0-3, 3-6)
- Pedals and steers a tricycle (Creative Curriculum); Riding a tricycle (HELP 0-3)
- Regulatory/Sensory organization (HELP 0-3)
- Child Reactivity/Distress under Indicator of Parent-Child Interaction (IGDI)
- Child Positive Engagement under Indicator of Parent-Child Interaction (IGDI)
- Sleep patterns and behaviors (HELP 0-3)

Specific Rules
A number of specific rules used to guide reoccurring issues and reasoning have been documented to provide better consistency across crosswalks. These rules include:

1. Assessment items about language used in the service of conversation are classified under Outcome 1.

2. Items about language skills without context or intent are under Outcome 2 (vocabulary, grammar, sentence structure, listening, and speaking).

3. Language skills and use related to getting needs meet are under Outcome 3.

4. Sense of self or self-concept skills are often distributed across the three outcomes. Items that are knowledge about self and distinctions from others are under outcome 2. Items focused on asserting oneself or conveying personal needs and desires to accomplish

---

[2] Crosswalks examined for this list include: HELP 0-3, HELP 3-6, ASQ, CC-Pre, Ounce, WS, High/Scope 0-3, High/Scope 3-6, BDI-2, Brigance IED-II, IGDI.

goals are under outcome 3. Items involving the self engaged with others are categorized under outcome 1. Responds to name is under outcome 1. But referring to self by name or telling first and last name, age, gender, address, or birthday are under Outcome 2. Plays with mirror image is outcome 1, but recognizing self and others in the mirror and distinguishing oneself in a photograph are outcome 2. Demonstrating self desires are often under outcome 3, including things like making choices between possibilities, saying no or resisting others' attempts to feed, identifying objects as "mine," asking for snacks, being selective about tasks, and showing persistence in choosing/continuing activities.

5. Learning skills without context are under Outcome 2.

6. Motor skills, either in isolation or involved with manipulating toys/objects, are under Outcome 3 and are asterisked (see General Rule 5).

7. Emotional expression and awareness (including showing pride in achievements or guilt/shame, etc.) are under Outcome 1.

8. Solving problems related to getting needs met is under Outcome 3 (e.g., experimenting with a brush to keep paint from dripping). Solving interpersonal problems is under Outcome 1. General problem solving (e.g., completes a 4 piece puzzle) is under Outcome 2.

9. Pre-literacy skills such as pre-writing and turning pages are double classified as both Outcomes 2 and 3 because of the contribution of the skill toward later literacy knowledge and skills as well as having elements of motor skills and taking actions to meet needs represented in Outcome 3. This double categorization includes things like copying or imitating forms (e.g., lines, crosses, circles). However, more specific draw-a-person and related body parts skills are only in outcome 2. Subsequent writing skills (e.g., prints personal data, prints letters in sequence, quality of printing) are only under Outcome 2. Cutting with scissors is only under Outcome 3.

10. Regulatory/sensory organization was not included in the crosswalk. This is a process underlying the achievement of all three outcome areas. Information about a child's regulatory/sensory organization does not provide unique knowledge about the child's functioning in a given outcome across settings and situations because it is seen as applying to all three outcomes.

11. Approaches to learning are typically categorized as Outcome 2 and Outcomes 3 (see Work Sampling).

12. Attention is usually categorized under Outcomes 2 and 3.

13. Imitation is usually under outcome 2, including imitation of household activities and varying levels of pretend and dramatic play.

Child Outcome Summary (COS) Process
Discussion Prompts

The pages that follow provide a few ideas for some types of questions or prompts that could be used to elicit conversation about a child's functioning with regard to the three global child outcome statements. As teams discuss child functioning in these outcomes areas, they generally draw on many sources of information and ask excellent questions that provide a specific description of what the child generally does with regard to each outcome. However, some teams have looked for further guidance about the kinds of questions that might help them focus on functional skills and span many of the components reflected in each outcome. The list that follows is by no means a comprehensive list of the types of questions or topics that might be discussed. It also is not intended to be used as a checklist necessary for discussion or as a checklist that will always constitute a complete discussion. However, it might provide some ideas to expand team approaches. It also may be helpful if individuals new to the COS process are quickly training other staff in using it and want more information for that purpose. As you begin to use this resource, we encourage you to share comments and additions with us at staff@the-eco-center.org so that we can include and circulate them as well!

Outcome 1: Child has positive social relationships.

Thinking about relating to adults, relating to other children, and (for those older than 18 months) following rules related to groups or interacting with others.

- Δ How does the child relate to his/her parent(s)?
- Δ How does the child relate to other relatives or extended family and close family friends (e.g., grandparents, aunts, extended kin, etc.)? Do these interactions with people differ depending on the setting the child is in with these people?
- Δ How does the child interact with familiar caregivers (e.g., child care providers, babysitters)?
- Δ How does the child relate to strangers? At first? After a while? In different settings and using different approaches?
- Δ How does the child interact with/respond to people in community settings (e.g., park, library, church, grocery store, with neighbors on walks, at the bus stop, in restaurants, at playgroups or outings, etc.)?
- Δ How does the child interact with/react to peers (e.g., at child care, in the park, in the neighborhood, in brief interactions in stores or at restaurants)?
- Δ How does the child relate to his/her siblings, cousins, or kids he/she sees frequently?
- Δ What is the child's eye contact with others like? Does it differ across situations or with different people?
- Δ How does the child display his/her emotions?
- Δ How does the child read and react to the emotions and expressions of others?
- Δ How does the child respond to touch from others?
- Δ How does the child maintain interactions with people?
- Δ In what situations and ways does the child express delight or display affection?
- Δ In the child's interactions, are there behaviors that may interfere with relationships or seem inappropriate in interactions expected for the child's age (e.g., screaming, biting, tantrums)? How often does this occur? In what situations? In what situations does it not occur?
- Δ Does the child display awareness of routines? How?
- Δ How does the child respond to transitions in routines or activities? Are the child's actions different for familiar transitions versus new transitions, or different across settings or with different people?
- Δ How and in what situations are interactions with others initiated?
- Δ How does the child engage in mutual activity (e.g., joint attention, communicate to convey desire to engage, initiate interaction or play, follow rules for mutual games)?
- Δ Does the child seek out others after an accomplishment? How?
- Δ Does the child seek out others after frustration or when angry? How?
- Δ Does the child participate in games (e.g., social, cooperative, rule-based, with turn-taking)? What do the child's interactions look like in these situations?
- Δ Does the child display an awareness of rules and expectations? How? Does the child behave differently in different contexts (e.g., quieter in church, more active outside)?
- Δ Does the child attempt to resolve his/her conflicts? How? What do these actions look like with peers, parents, etc.?
- Δ How does the child respond when others are not attending to him/her?
- Δ How does the child respond when someone arrives? Someone new? Someone familiar? How does the child respond when someone leaves?
- Δ Talk about the child's functioning with regard to turn-taking, showing, and sharing? With adults? With other children?

2

The Early Childhood Outcomes Center   **11-8-12**

Δ How would you expect other children this age to act in these situations?

Outcome 2: Child acquires and uses knowledge and skills.

Thinking, reasoning, remembering, and problem solving; understanding symbols; and understanding the physical and social worlds.

- Δ How does the child use the words and skills she/he has in everyday settings (e.g., at home, at the park, at child care, at the store, with other kids, at child care, in restaurants, with different people)?
- Δ Tell me about a time when he/she tried to solve a problem (e.g., overcome an obstacle/problem interfering with something important to him/her). What did he/she do?
- Δ What concepts does the child understand? Does the child incorporate these into strategies that he/she uses to accomplish something meaningful? How?
- Δ How does the child understand and respond to directions and requests from others?
- Δ How does the child imitate others' actions (e.g., peers, adults) across settings to learn or try new things?
- Δ How does the child display understanding of differences in roles, characteristics, and expectations across people and situations (with increasing age role understanding may change from immediate household roles and differences to more external community helper roles)?
- Δ Can the child use his/her understanding to communicate problems or attempt the solutions that others suggest (e.g., try new strategies that they haven't thought of based on gestures or suggestions using words they know)?
- Δ Can the child answer questions of interest in meaningful ways?
- Δ Does the child use something learned at one time at a later time or in another situation?
- Δ Does the child display an awareness of the distinctions between things (e.g., object characteristics, size differences, differences in object functions)?
- Δ What does the child do if an action or a strategy attempted isn't successful? (e.g., how does he/she try to modify approach, show persistence, etc.)
- Δ How does the child demonstrate her/his understanding of symbols into concepts, communication, and play?
- Δ How does the child interact with books, pictures, and print?
- Δ How does the child's play suggest understanding of familiar scripts for how things work, what things are related, what comes next, and memory of previous actions in that situation?
- Δ Does the child's play show attempts to modify strategies/approaches and to try new things? How?
- Δ Are there kinds of knowledge and skills that are not similar to same age peers and/or that might interfere with acquiring and using knowledge and skills?

Δ How would you expect other children this age to act in these situations?

The Early Childhood Outcomes Center   **11-8-12**

Outcome 3: Child takes appropriate action to meet his/her needs.

Taking care of basic needs; getting from place to place and using tools; and (if older than 24 months) contributing to own health and safety.

Δ What does the child do when she/he can't get or doesn't have what she wants?

Δ What does the child do when he/she wants something that is out of reach or hard to get?

Δ What does the child do when he/she is upset or needs comfort?

Δ What does the child do when she/he is hungry?

Δ What does he/she do when he/she is frustrated?

Δ What does the child do when she/he needs help?

Δ How does the child convey his/her needs?

Δ How are the child's actions to seek help or to convey his/her needs different from one setting to another? How do they differ with different people? (e.g., child care vs. home vs. community setting, with parent vs. grandparent, familiar person vs stranger)

Δ Tell me about the child's actions when dressing and/or undressing?

Δ What does the child do before and after peeing and pooping?

Δ What does the child do at mealtime (eating, drinking)? Are there differences across settings and with different people?

Δ How does the child get started playing with toys? What does the child do when he/she is interested in a different toy than he/she has?

Δ Tell me about the child's actions/reactions with regard to hygiene (toothbrushing, washing hands/face, blowing nose, etc.)?

Δ Does the child show awareness of situations that might be dangerous? What does he/she do (give examples, (e.g., to dropoffs, hot stoves, cars/crossing streets, strangers, etc.)?

Δ Are there situations when a problem behavior or disability interferes with the child's ability to take action to meet needs? How consistently? How serious is it? Does the child take alternative approaches? What are those?

Δ Are the actions the child uses to meet his/her needs appropriate for his/her age? Can he/she accomplish the things that peers do?

Δ How does the child respond to delays in receiving expected attention and/or help from others?

Δ How does the child respond to challenges?

Δ Does the child display toy preferences? How do you know?

Δ How does the child get from place to place when desired or needed?

Δ What does the child do when she/he is bored? How does she/he amuse her/himself or seek out something fun?

Δ How does the child respond to problematic or unwanted peer behavior?

Δ How does the child use materials to have an effect (e.g., drawing materials, tools, etc.)?

Δ **How would you expect other children this age to act in these situations?**

Appendix C

State Requirement Materials from the U.S. Department of Education

2015 Part C APR Memo

Part C SPP/APR Indicator/Measurement Table

2015 Part B APR Memo

Part B SPP/APR Indicator/Measurement Table

For additional information:
https://osep.grads360.org/#program/spp-apr-resources

DEC 2 3 2014

| **Contact Person** |
| Name: Gregg Corr |
| Telephone: 202-245-7309 |

OSEP 15-05

MEMORANDUM

TO: Lead Agency Directors, Part C Coordinators, State Interagency Coordinating Council Chairpersons, and State Data Managers

FROM: Melody Musgrove, Ed.D.
Director
Office of Special Education Programs (OSEP)

SUBJECT: Part C State Performance Plan (Part C – SPP) and Part C Annual Performance Report (Part C – APR)

**ACTION
REQUIRED:** Submission of the Part C – SPP and the Part C – APR by February 2, 2015

In accordance with 20 U.S.C. 1416(b)(1)(C) and 1442 of the Individuals with Disabilities Education Act (IDEA), each Lead Agency must amend, at least once every six years, its Part C State Performance Plan (Part C – SPP), and under 20 U.S.C. 1416(b)(2)(C)(ii)(II) and 1442, report annually, through the Part C Annual Performance Report (Part C – APR) to the Secretary on the State's performance under its Part C – SPP. The Part C – SPP evaluates the State's efforts to implement the requirements and purposes of Part C and describes how the State will improve such implementation. The original Part C – SPP that all States submitted in 2005 covered a period of six years for Federal fiscal years (FFYs) 2005 through 2010 and OSEP extended the Part C – SPP for two years to cover FFYs 2011 and 2012.

On February 2, 2015, each Lead agency must submit a new Part C – SPP that covers the six year period for FFYs 2013 through 2018 and includes a new Indicator 11, the State Systemic Improvement Plan (SSIP) that is part of OSEP's Results Driven Accountability Framework (RDA). In accordance with 20 U.S.C. 1416(b)(2)(C)(ii)(I) and 1442 and 34 CFR §303.702(b)(1)(i)(A), each Lead Agency must also report annually to the public on the performance of each early intervention service (EIS) program located in the State on the targets in its Part C – SPP, as soon as practicable, but no later than 120 days following the State's submission of its Part C – APR to the Secretary.

Information Collection 1820-0578, *Part C State Performance Plan (SPP)/Annual Performance Report (APR)*, is available electronically at https://osep.grads360.org/#program/spp-apr-resources and contains both Part C – SPP and Part C – APR instructions, and the document is combined as one report and referred to throughout this memo as the SPP/APR. The link also includes the *Part C Related Requirements* document as a reference. Beginning with the FFY 2013 SPP/APR, the State must submit its SPP/APR online using the SPP/APR module on GRADS 360° (https://osep.grads360.org). The online SPP/APR module is only accessible to State-authorized users who have a unique log-in user name and password. User names and passwords are available to all State-authorized users by emailing EDEN_ss@ed.gov. States are encouraged to make note of the due dates and implement a development schedule accordingly.

By February 2, 2015 (unless specifically indicated otherwise below), States must submit:

1) One report that includes both SPP and APR information.

2) An introduction, with sufficient detail so that the Secretary and the public are informed of and understand the State's systems designed to drive improved results for infants and toddlers with disabilities and to ensure that the Lead Agency meets the requirements of IDEA Part C. This introduction must include descriptions of the State's:

   a. General Supervision System: The systems that are in place for the State Lead Agency, using a single line of responsibility, to ensure that IDEA Part C requirements are met, e.g., monitoring of EIS programs and providers, dispute resolution, etc.;

   b. Technical Assistance System: The mechanisms that the State has in place to ensure the timely delivery of high quality, evidenced based technical assistance and support to EIS programs;

   c. Professional Development System: The mechanisms the State has in place to ensure that EIS providers have the skills to effectively provide services that improve results for infants and toddlers with disabilities and their families;

   d. Stakeholder Involvement: The mechanism for soliciting broad stakeholder input on the State's targets in the SPP/APR and the development and implementation of new Indicator 11, the SSIP (Note that all of the information regarding Indicator 11, including the baseline data for FFY 2013, targets for FFYs 2014 through 2018, and other content is not due until April 1, 2015, as reflected in items 3, 4 and 7 below)[1]; and

   e. Reporting to the Public: How and where the State reported to the public on the FFY 2012 performance of each EIS program located in the State on the targets in the SPP/APR as soon as practicable, but not later than 120 days following the State's submission of its FFY 2012 APR, as required by 34 CFR §303.702(b)(1)(i)(A); and a description of where, on its Web site, a complete copy of the State's SPP, including any revision if the State has revised the SPP that it submitted with its FFY 2012 APR in 2014, is available.

---

[1] As noted in the conference report to HR 1350, it is Congress' expectation that targets will be developed with broad stakeholder input and will be disseminated to the public.

3) Baseline data for FFY 2013 for Indicator 11.[2]

4) Annual State targets for Indicators 1 through 10, determined with stakeholder input, for each year and that cover the years of the SPP (i.e., FFY 2013 through FFY 2018). Targets for Indicator 11, which cover FFY 2014 through FFY 2018.

5) Data from FFY 2013 and other responsive APR information for Indicators 1 through 10.[3]

6) An explanation of any slippage in indicators where the State did not meet its FFY 2013 target.

7) The specific content required to complete Phase I of the SSIP required by Indicator 11. Phase I includes a detailed analysis that will guide the selection of coherent improvement strategies to increase the State capacity to lead meaningful change in EIS programs related to the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families. (See Indicator 11 for specific content of Phases II and III of the SSIP, which must be included with the States' FFY 2014 through FFY 2018 SPP/APRs.):

   a. Data Analysis

   b. Analysis of State Infrastructure to Support Improvement and Build Capacity

   c. State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families

   d. Selection of Coherent Improvement Strategies

   e. Theory of Action

8) Information to address any required actions identified in OSEP's letter responding to the State's February 3, 2014 submission of its FFY 2012 APR, including providing information in an attachment to address required actions identified in previous Indicator 9 (Timely Correction).

When completing the SPP and APR, Lead Agencies will need to use the following parts of Information Collections 1820-0578:

- *SPP/APR Instructions*

- *Part C Indicator Measurement Table* with Instructions

The *Part C Indicator Measurement Table* lists the Monitoring Priorities and Indicators, required data sources and measurement, and instructions for providing the required information for each indicator. In addition to the percentages required in the indicators, Lead Agencies are required to provide <u>actual numbers</u> used in the calculations.[4]

---

[2] For the FFY 2013 SPP/APR, all information related to Indicator 11, the SSIP, including baseline and targets, and specific content, is due on April 1, 2015.

[3] Previous Indicator 9 (Timely Correction) and Indicator 14 (Timely and Accurate Data) have been removed. Consequently, the indicators have been renumbered.

[4] Lead Agencies are not required to provide the actual numbers for Indicators 2, 5, 6, 9, and 10 because they are using 618 State-reported data for these Indicators.

The Department will review the information provided in the State's FFY 2013 APR, other State-reported data, information obtained through monitoring visits, and other public information and will determine, under IDEA sections 616(d) and 642, if the State meets the requirements and purposes of Part C of the IDEA or needs assistance, needs intervention, or needs substantial intervention in implementing the requirements of Part C of the IDEA. When reporting on correction of noncompliance in the APR, each State must include confirmation that: (1) each EIS program has corrected each individual instance of child-specific noncompliance; and (2) the State has verified that the EIS program is currently correctly implementing the specific regulatory requirement. The State should review IDEA section 616(e) regarding the potential enforcement actions the Department is required to take as a result of, and the potential future impact of, the Department's annual determination.

Prior to issuing the Department's 2015 determination for each State, OSEP will offer the State the opportunity to clarify or correct the data submitted in its FFY 2013 SPP/APR. OSEP will provide feedback through the SPP/APR module on the State's February 2, 2015 FFY 2013 APR data that OSEP will consider in making the Department's determination under IDEA sections 616(d) and 642. In response to OSEP's feedback, the State must submit to OSEP through the SPP/APR module any corrected data and clarify any misunderstandings by OSEP about the data submitted. In order for the State's clarifications or corrections to data to be considered, the State's response must be certified as completed no later than close-of-business on the date specified in OSEP's SPP/APR module summary.

Please note that any State that does not meet the February 2, 2015[5] timeline for submitting data for each indicator will not be permitted to submit clarifications or corrections for purposes of the Department's determination under IDEA sections 616(d) and 642.

In accordance with sections 616(e) and 642 of the IDEA, in the Department's June 23, 2014 determination letters, the Secretary advised States that were needs assistance for two or more consecutive years of available sources of technical assistance. The Secretary directed such a State to determine the result(s) and/or compliance indicator(s), and improvement strategies, on which it will focus its use of available technical assistance in order to improve its performance. The Secretary strongly encouraged each State to access technical assistance related to those specific SPP/APR compliance indicator(s) for which the State received a score of zero on the 2014 Compliance Matrix. The State must report with its FFY 2013 APR submission (due February 2, 2015) on: (1) the technical assistance sources from which the State received assistance; and (2) what actions the State took as a result of that technical assistance. The extent to which your State takes advantage of available technical assistance may affect the actions we take under IDEA sections 616 and 642, should your State not be identified as meets requirements in 2015.

If you have any further questions about the Part C - SPP and/or Part C - APR and/or the submission options listed above, please contact your OSEP Part C State Contact.

Enclosures

cc: National Center for Systemic Improvement (NCSI)
    Early Childhood Technical Assistance (ECTA) Center

---

[5] See Note 2

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
Monitoring Priority: Early Intervention Services In Natural Environments		
1. Percent of infants and toddlers with IFSPs who receive the early intervention services on their IFSPs in a timely manner. (20 U.S.C. 1416(a)(3)(A) and 1442)	**Data Source:** Data to be taken from monitoring or State data system and must be based on actual, not an average, number of days. Include the State's criteria for "timely" receipt of early intervention services, i.e., the time period from parent consent to when IFSP services are actually initiated. **Measurement:** Percent = [(# of infants and toddlers with IFSPs who receive the early intervention services on their IFSPs in a timely manner) divided by the (total # of infants and toddlers with IFSPs)] times 100. Account for untimely receipt of services, including the reasons for delays.	*If data are from State monitoring, describe the method used to select EIS programs for monitoring. If data are from a State database, describe the time period in which the data were collected (e.g., September through December, fourth quarter, selection from the full reporting period) and how the data accurately reflect data for infants and toddlers with IFSPs for the full reporting period.* Targets must be 100%. Describe the results of the calculations and compare the results to the target. Describe the method used to collect these data and if data are from the State's monitoring, describe the procedures used to collect these data. States report in both the numerator and denominator under Indicator 1 on the number of children for whom the State ensured the timely initiation of new services identified on the IFSP. Include the timely initiation of new early intervention services from both initial IFSPs and subsequent IFSPs. Provide actual numbers used in the calculation. The State's timeliness measure for this indicator must be either: (1) a time period that runs from when the parent consents to IFSP services; or (2) the IFSP initiation date (established by the IFSP Team, including the parent). States are not required to report in their calculation the number of children for whom the State has

---

[1] Monitoring Priorities, indicators, and measurements included on the *Part C Indicator Measurement Table* are to be used to populate designated sections of the SPP and APR Templates. Populated templates can be found at http://spp-apr-calendar.rrfcnetwork.org/explorer/view/id/446/?3#category3

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
		identified the cause for the delay as exceptional family circumstances documented in the child's record. If a State chooses to report in its calculation children for whom the State has identified the cause for the delay as exceptional family circumstances documented in the child's record, the numbers of these children are to be included in the numerator and denominator. Include in the discussion of the data, the numbers the State used to determine its calculation under this indicator and report separately the number of documented delays attributable to exceptional family circumstances. Provide detailed information about the timely correction of noncompliance as noted in OSEP's response table for the previous APR. If the State did not ensure timely correction of the previous noncompliance, provide information on the extent to which noncompliance was subsequently corrected (more than one year after identification). In addition, provide information regarding the nature of any continuing noncompliance, methods to ensure correction, and any enforcement actions that were taken.
2. Percent of infants and toddlers with IFSPs who primarily receive early intervention services in the home or community-based settings. (20 U.S.C. 1416(a)(3)(A) and 1442)	**Data Source:** Data collected under IDEA section 618 **Measurement:** Percent = [(# of infants and toddlers with IFSPs who primarily receive early intervention services in the home or community-based settings) divided by the (total # of infants and toddlers with IFSPs)] times 100.	*For this indicator, report 618 data that were collected on a date between October 1 and December 1, 2013 and due on February 1, 2014. Sampling from State's 618 data is not allowed.* Describe the results of the calculations and compare the results to the target. The data reported in this indicator should be consistent with the State's reported 618 data reported in Table 2. If not, explain.
3. Percent of infants and toddlers with IFSPs who demonstrate improved: A. Positive social-emotional skills	**Data Source:** State selected data source.	*Sampling of **infants and toddlers with IFSPs** is allowed. When sampling is used, submit a description of the sampling methodology outlining how the design will yield valid and reliable estimates.*

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
(including social relationships); B. Acquisition and use of knowledge and skills (including early language/communication); and C. Use of appropriate behaviors to meet their needs. (20 U.S.C. 1416(a)(3)(A) and 1442)	**Measurement:** Outcomes: A. Positive social-emotional skills (including social relationships); B. Acquisition and use of knowledge and skills (including early language/communication); and C. Use of appropriate behaviors to meet their needs. Progress categories for A, B and C: a. Percent of infants and toddlers who did not improve functioning = [(# of infants and toddlers who did not improve functioning) divided by (# of infants and toddlers with IFSPs assessed)] times 100. b. Percent of infants and toddlers who improved functioning but not sufficient to move nearer to functioning comparable to same-aged peers = [(# of infants and toddlers who improved functioning but not sufficient to move nearer to functioning comparable to same-aged peers) divided by (# of infants and toddlers with IFSPs assessed)] times 100. c. Percent of infants and toddlers who improved functioning to a level nearer to same-aged peers but did not reach it = [(# of infants and toddlers who improved functioning to a level nearer to same-aged peers but did not reach it) divided by (# of infants and toddlers with IFSPs assessed)] times 100. d. Percent of infants and toddlers who improved functioning to reach a level comparable to same-aged peers = [(# of infants and toddlers who improved functioning to reach a level comparable to same-aged peers) divided by	*(See General Instructions page 2 for additional instructions on sampling.)* Describe the results of the calculations and compare the results to the targets. States will use the progress categories for each of the three Outcomes to calculate and report the two Summary Statements. Report progress data and calculate Summary Statements to compare against the six targets. Provide the actual numbers and percentages for the five reporting categories for each of the three outcomes. In presenting results, provide the criteria for defining "comparable to same-aged peers." If a State is using the Early Childhood Outcomes Center (ECO) Child Outcomes Summary Form (COSF), then the criteria for defining "comparable to same-aged peers" has been defined as a child who has been assigned a scored of 6 or 7 on the COSF. In addition, list the instruments and procedures used to gather data for this indicator, including if the State is using the ECO COSF. If the State's Part C eligibility criteria include infants and toddlers who are at risk of having substantial developmental delays (or "at-risk infants and toddlers") under IDEA section 632(5)(B)(i), the State must report data in two ways. First, it must report on all eligible children but exclude its at-risk infants and toddlers (i.e., include just those infants and toddlers experiencing developmental delay (or "developmentally delayed children") or having a diagnosed physical or mental condition that has a high probability of resulting in developmental delay (or "children with diagnosed conditions"). Second, the State must separately report outcome data on either: (1) just its at-risk infants and toddlers; or (2) aggregated performance data on all of the infants

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
	(# of infants and toddlers with IFSPs assessed)] times 100. e. Percent of infants and toddlers who maintained functioning at a level comparable to same-aged peers = [(# of infants and toddlers who maintained functioning at a level comparable to same-aged peers) divided by (# of infants and toddlers with IFSPs assessed)] times 100. **Summary Statements for Each of the Three Outcomes:** **Summary Statement 1:** Of those infants and toddlers who entered or exited early intervention below age expectations in each Outcome, the percent who substantially increased their rate of growth by the time they turned 3 years of age or exited the program. **Measurement for Summary Statement 1:** Percent = # of infants and toddlers reported in progress category (c) plus # of infants and toddlers reported in category (d) divided by [# of infants and toddlers reported in progress category (a) plus # of infants and toddlers reported in progress category (b) plus # of infants and toddlers reported in progress category (c) plus # of infants and toddlers reported in progress category (d)] times 100. **Summary Statement 2:** The percent of infants and toddlers who were functioning within age expectations in each Outcome by the time they turned 3 years of age or exited the program. **Measurement for Summary Statement 2:** Percent = # of infants and toddlers reported in progress category (d) plus [# of infants and toddlers reported in progress category (e) divided by the total # of infants and toddlers reported in progress	and toddlers it serves under Part C (including developmentally delayed children, children with diagnosed conditions, and at-risk infants and toddlers).

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
	categories (a) + (b) + (c) + (d) + (e)] times 100.	
4. Percent of families participating in Part C who report that early intervention services have helped the family: A. Know their rights; B. Effectively communicate their children's needs; and C. Help their children develop and learn. (20 U.S.C. 1416(a)(3)(A) and 1442)	**Data Source:** State selected data source. State must clarify the data source in the State Performance Plan. **Measurement:** A. Percent = [(# of respondent families participating in Part C who report that early intervention services have helped the family know their rights) divided by the (# of respondent families participating in Part C)] times 100. B. Percent = [(# of respondent families participating in Part C who report that early intervention services have helped the family effectively communicate their children's needs) divided by the (# of respondent families participating in Part C)] times 100. C. Percent = [(# of respondent families participating in Part C who report that early intervention services have helped the family help their children develop and learn) divided by the (# of respondent families participating in Part C)] times 100.	*Sampling of **families participating in Part C** is allowed. When sampling is used, a description of the sampling methodology outlining how the design will yield valid and reliable estimates must be submitted to OSEP. (See General Instructions page 2 for additional instruction on sampling.)* States should describe the results of the calculations and compare the results to the target. Include a description of how the State has ensured that any response data are valid and reliable, including how the data represent the demographics of the State. Provide the actual numbers used in the calculation. If States are using a survey and the survey is revised or a new survey is adopted, States must submit a copy with the APR.
Monitoring Priority: Effective General Supervision Part C		
Effective General Supervision Part C / Child Find		
5. Percent of infants and toddlers birth to 1 with IFSPs compared to national data. (20 U.S.C. 1416(a)(3)(B) and 1442)	**Data Source:** Data collected under IDEA section 618 **Measurement:** Percent=[(# of infants and toddlers birth to 1 with IFSPs) divided by the (population of infants and toddlers birth to 1)] times 100 compared to national	*For this indicator, report 618 data that were collected on a date between October 1 and December 1, 2013 and due on February 1, 2014. Sampling from State's 618 data is not allowed.* Describe the results of the calculations and compare the results to the target and to national data. The data reported in this indicator should be consistent with the State's reported 618 data reported in Table

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
	data.	1. If not, explain.
6. Percent of infants and toddlers birth to 3 with IFSPs compared to national data. (20 U.S.C. 1416(a)(3)(B) and 1442)	**Data Source:** Data collected under IDEA section 618 **Measurement:** Percent=[(# of infants and toddlers birth to 3 with IFSPs) divided by the (population of infants and toddlers birth to 3)] times 100 compared to national data.	*For this indicator, report 618 data that were collected on a date between October 1 and December 1, 2013 and due on February 1, 2014. Sampling from State's 618 data is not allowed.* Describe the results of the calculations and compare the results to the target and to national data. The data reported in this indicator should be consistent with the State's reported 618 data reported in Table 1. If not, explain.
7. Percent of eligible infants and toddlers with IFSPs for whom an initial evaluation and initial assessment and an initial IFSP meeting were conducted within Part C's 45-day timeline. (20 U.S.C. 1416(a)(3)(B) and 1442)	**Data Source:** Data to be taken from monitoring or State data system and must address the timeline from point of referral to initial IFSP meeting based on actual, not an average, number of days. **Measurement:** Percent = [(# of eligible infants and toddlers with IFSPs for whom an initial evaluation and initial assessment and an initial IFSP meeting were conducted within Part C's 45-day timeline) divided by the (# of eligible infants and toddlers evaluated and assessed for whom an initial IFSP meeting was required to be conducted)] times 100. Account for untimely evaluations, assessments, and initial IFSP meetings, including the reasons for delays.	*If data are from State monitoring, describe the method used to select EIS programs for monitoring. If data are from a State database, describe the time period in which the data were collected (e.g., September through December, fourth quarter, selection from the full reporting period) and how the data accurately reflect data for infants and toddlers with IFSPs for the full reporting period.* Targets must be 100%. Describe the results of the calculations and compare the results to the target. Describe the method used to collect these data and if data are from the State's monitoring, describe the procedures used to collect these data. Provide actual numbers used in the calculation. States are not required to report in their calculation the number of children for whom the State has identified the cause for the delay as exceptional family circumstances, as defined in 34 CFR §303.310(b), documented in the child's record. If a State chooses to report in its calculation children for whom the State has identified the cause for the delay as exceptional family circumstances documented in the child's record, the numbers of these children are to be included in the numerator and denominator.

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
		Include in the discussion of the data, the numbers the State used to determine its calculation under this indicator and report separately the number of documented delays attributable to exceptional family circumstances.

Provide detailed information about the timely correction of noncompliance as noted in OSEP's response table for the previous APR. If the State did not ensure timely correction of the previous noncompliance, provide information on the extent to which noncompliance was subsequently corrected (more than one year after identification). In addition, provide information regarding the nature of any continuing noncompliance, methods to ensure correction, and any enforcement actions that were taken. |
| **Effective General Supervision Part C / Effective Transition** | | |
| 8. The percentage of toddlers with disabilities exiting Part C with timely transition planning for whom the Lead Agency has:

A. Developed an IFSP with transition steps and services at least 90 days, and at the discretion of all parties, not more than nine months, prior to the toddler's third birthday;

B. Notified (consistent with any opt-out policy adopted by the State) the SEA and the LEA where the toddler resides at least 90 days prior to the toddler's third birthday for toddlers potentially eligible for Part B preschool services; and

C. Conducted the transition conference held with the approval of the family at least 90 days, and at the discretion of | **Data Source:**

Data to be taken from monitoring or State data system.

Measurement:

A. Percent = [(# of toddlers with disabilities exiting Part C who have an IFSP with transition steps and services at least 90 days, and at the discretion of all parties not more than nine months, prior to their third birthday) divided by the (# of toddlers with disabilities exiting Part C)] times 100.

B. Percent = [(# of toddlers with disabilities exiting Part C where notification (consistent with any opt-out policy adopted by the State) to the SEA and LEA occurred at least 90 days prior to their third birthday for toddlers potentially eligible for Part B preschool services) divided by the (# of toddlers | Indicators 8A, 8B, and 8C: Targets must be 100%.

Describe the results of the calculations and compare the results to the target. Describe the method used to collect these data. Provide the actual numbers used in the calculation.

Indicators 8A and 8C: If data are from the State's monitoring, describe the procedures used to collect these data. If data are from State monitoring, also describe the method used to select EIS programs for monitoring. If data are from a State database, describe the time period in which the data were collected (e.g., September through December, fourth quarter, selection from the full reporting period) and how the data accurately reflect data for infants and toddlers with IFSPs for the full reporting period.

Indicator 8A: States are not required to report in their calculation the number of children for whom the State has identified the cause for the delay as |

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
all parties, not more than nine months, prior to the toddler's third birthday for toddlers potentially eligible for Part B preschool services. (20 U.S.C. 1416(a)(3)(B) and 1442)	with disabilities exiting Part C who were potentially eligible for Part B)] times 100. C. Percent = [(# of toddlers with disabilities exiting Part C where the transition conference occurred at least 90 days, and at the discretion of all parties at least nine months, prior to the toddler's third birthday for toddlers potentially eligible for Part B) divided by the (# of toddlers with disabilities exiting Part C who were potentially eligible for Part B)] times 100. Account for untimely transition planning under 8A, 8B, and 8C, including the reasons for delays.	exceptional family circumstances documented in the child's record. If a State chooses to report in its calculation children for whom the State has identified the cause for the delay as exceptional family circumstances documented in the child's record, the numbers of these children are to be included in the numerator and denominator. Include in the discussion of the data, the numbers the State used to determine its calculation under this indicator and report separately the number of documented delays attributable to exceptional family circumstances. Indicator 8B: Under 34 CFR §303.401(e), the State may adopt a written policy that requires the lead agency to provide notice to the parent of an eligible child with an IFSP of the impending notification to the SEA and LEA under IDEA section 637(a)(9)(A)(ii)(I) and 34 CFR §303.209(b)(1) and (2) and permits the parent within a specified time period to "opt-out" of the referral. Under the State's opt-out policy, the State is not required to include in the calculation under 8B (in either the numerator or denominator) the number of children for whom the parents have opted out. However, the State must include in the discussion of data, the number of parents who opted out. In addition, any written opt-out policy must be on file with the Department as part of the State's Part C application under IDEA section 637(a)(9)(A)(ii)(I) and 34 CFR §§303.209(b) and 303.401(d). Indicator 8C: Do not include in the calculation, but provide a separate number for those toddlers for whom the parent did not provide approval for the transition conference. Indicator 8C: States are not required to report in their calculation the number of children for whom the State has identified the cause for the delay as exceptional family circumstances documented in the child's record. If a State chooses to report in its calculation children for whom the State has identified

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
		the cause for the delay as exceptional family circumstances documented in the child's record, the numbers of these children are to be included in the numerator and denominator. Include in the discussion of the data, the numbers the State used to determine its calculation under this indicator and report separately the number of documented delays attributable to exceptional family circumstances.
		Indicators 8A, 8B, and 8C: Provide detailed information about the timely correction of noncompliance as noted in OSEP's response table for the previous APR. If the State did not ensure timely correction of the previous noncompliance, provide information on the extent to which noncompliance was subsequently corrected (more than one year after identification). In addition, provide information regarding the nature of any continuing noncompliance, methods to ensure correction, and any enforcement actions that were taken.
Effective General Supervision Part C / General Supervision		
9. Percent of hearing requests that went to resolution sessions that were resolved through resolution session settlement agreements (applicable if Part B due process procedures are adopted). (20 U.S.C. 1416(a)(3)(B) and 1442)	**Data Source:** Data collected under IDEA section 618 **Measurement:** Percent = (3.1(a) divided by 3.1) times 100.	*Sampling is not allowed.* Describe the results of the calculations and compare the results to the target. States are not required to establish baseline or targets if the number of resolution sessions is less than 10. In a reporting period when the number of resolution sessions reaches ten or greater, the State must develop baseline and targets and report them in the corresponding APR. States may express their targets in a range, e.g., 75-85%. If the data reported in this indicator are not the same as the State's 618 data, explain. States are not required to report data at the EIS

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
		program level.
10. Percent of mediations held that resulted in mediation agreements. (20 U.S.C. 1416(a)(3)(B) and 1442)	**Data Source:** Data collected under IDEA section 618 **Measurement:** Percent = [(2.1(a)(i) + 2.1(b)(i)) divided by 2.1] times 100.	*Sampling is not allowed.* Describe the results of the calculations and compare the results to the target. States are not required to establish baseline or targets if the number of mediations is less than 10. In a reporting period when the number of mediations reaches ten or greater, the State must develop baseline and report them in the corresponding APR. The consensus among mediation practitioners is that 75-85% is a reasonable rate of mediations that result in agreements and is consistent with national mediation success rate data. States may express their targets in a range, e.g., 75-85%. If the data reported in this indicator are not the same as the State's 618 data, explain. States are not required to report data at the EIS program level.

INDICATOR: The State's SPP/APR includes a State Systemic Improvement Plan (SSIP) that meets the requirements set forth for this indicator.

MEASUREMENT: The State's SPP/APR includes an SSIP that is a comprehensive, ambitious, yet achievable multi-year plan for improving results for infants and toddlers with disabilities and their families. The SSIP includes each of the components described below.

INSTRUCTIONS FOR THE INDICATOR/MEASUREMENT –

Baseline Data: In its FFY 2013 SPP/APR, due February 1, 2015, the State must provide FFY 2013 baseline data that must be expressed as a percentage and which is aligned with the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families.

Targets: In its FFY 2013 SPP/APR, due February 1, 2015, the State must provide measurable and rigorous targets (expressed as percentages) for each of the five years from FFY 2014 through-FFY 2018. The State's FFY 2018 target must demonstrate improvement over the State's FFY 2013 baseline data.

Updated data: In its FFYs 2014 through FFY 2018 SPPs/APRs, due February 2016 through February 2020, the State must provide updated data for that specific FFY (expressed as percentages) and that data must be aligned with the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families. In its FFYs 2014 through FFY 2018 SPPs/APRs, the State must report on whether it met its target.

OVERVIEW OF THE THREE PHASES OF THE SSIP: It is of the utmost importance to improve results for infants and toddlers with disabilities and their families by improving early intervention services. Stakeholders, including parents of infants and toddlers with disabilities, early intervention service (EIS) programs and providers, the State Interagency Coordinating Council, and others, are critical participants in improving results for infants and toddlers with disabilities and their families and must be included in developing, implementing, evaluating, and revising the SSIP and included in establishing the State's targets under Indicator 11. The SSIP should include information about stakeholder involvement in all three phases.

Phase I: Analysis (which the State must include with the February 2, 2015 submission of its SPP/APR for FFY 2013):

- Data Analysis;

- Analysis of State Infrastructure to Support Improvement and Build Capacity;

- State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families;

- Selection of Coherent Improvement Strategies; and

- Theory of Action.

Phase II: Plan (which, in addition to the Phase 1 content (including any updates) outlined above, the State must include with the February 1, 2016 submission of its SPP/APR for FFY 2014):

- Infrastructure Development;

- Support for EIS Program and/or EIS Provider Implementation of Evidence-Based Practices; and

- Evaluation.

Phase III: Implementation and Evaluation (which, in addition to the Phase I and Phase II content (including any updates) outlined above, the State must include with the February 1, 2017 submission of its SPP/APR for FFY 2015, and update in 2018, 2019, and 2020):

- Results of Ongoing Evaluation and Revisions to the SSIP.

SPECIFIC CONTENT OF EACH PHASE OF THE SSIP

Phase I: Analysis

Phase I of the SSIP includes a detailed analysis that will guide the selection of coherent improvement strategies to increase the State's capacity to lead meaningful change in EIS programs and/or EIS providers to improve results for infants and toddlers with disabilities and their families. Phase I must include the following five areas:

- Data Analysis: A description of how the State identified and analyzed key data, including data from SPP/APR indicators, 618 data collections, and other available data as applicable, to: (1) select the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families, and (2) identify root causes contributing to low performance. The description must include information about how the data were disaggregated by multiple variables (e.g., EIS program and/or EIS provider, geographic region, race/ethnicity, socioeconomic status, gender, etc.) As part of its data analysis, the State should also consider compliance data and whether those data present potential barriers to improvement. In addition, if the State identifies any concerns about the quality of the data, the description must include how the State will address these concerns. Finally, if additional data are needed, the description should include the methods and timelines to collect and analyze the additional data.

- Analysis of State Infrastructure to Support Improvement and Build Capacity: A description of how the State analyzed the capacity of its current infrastructure to support improvement and build capacity in EIS programs and/or EIS providers to implement, scale up, and sustain the use of evidence-based practices to improve results for infants and toddlers with disabilities and their families. State systems that make up its infrastructure include, at a minimum: governance, fiscal, quality standards, professional development, data, technical assistance, and accountability/monitoring. The description must include current strengths of the systems, the extent the systems are coordinated, and areas for improvement of functioning within and across the systems. The State must also identify current State-level improvement plans and other early learning initiatives, such as Race to the Top-Early Learning Challenge and the Home Visiting program and describe the extent that these new initiatives are aligned, and how they are, or could be, integrated with, the SSIP. Finally, the State should identify representatives (e.g., offices, agencies, positions, individuals, and other stakeholders) that were involved in developing Phase I of the SSIP and that will be involved in developing and implementing Phase II of the SSIP.

- State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families: A statement of the result(s) the State intends to achieve through the implementation of the SSIP. The State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families must be aligned to an SPP/APR indicator or a component of an SPP/APR indicator. The State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families must be clearly based on the Data and State Infrastructure Analyses and must be a child- or family-level outcome in contrast to a process outcome. The State may select a single result (e.g., increase the rate of growth in infants and toddlers demonstrating positive social-emotional skills) or a cluster of related results (e.g., increase the percentage reported under child outcome B under Indicator 3 of the SPP/APR (knowledge and skills) and increase the percentage trend reported for families under Indicator 4 (helping their child develop and learn)).

- Selection of Coherent Improvement Strategies: An explanation of how the improvement strategies were selected, and why they are sound, logical and aligned, and will lead to a measurable improvement in the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families. The improvement strategies should include the strategies, identified through the Data and State Infrastructure Analyses, that are needed to improve the State infrastructure and to support EIS program and/or EIS provider implementation of evidence-based practices to improve the State-identified result(s) for infants and toddlers with disabilities and their families. The State must describe how implementation of the improvement strategies will address identified root causes for low performance and ultimately build EIS program and/or EIS provider capacity to achieve the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families.

- Theory of Action: A graphic illustration that shows the rationale of how implementing the coherent set of improvement strategies selected will increase the State's capacity to lead meaningful change in EIS programs and/or EIS providers, and achieve improvement in the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families.

Phase II: Plan

The focus of Phase II is on building State capacity to support EIS programs and/or EIS providers with the implementation of evidence-based practices that will lead to measurable improvement in the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families. Phase II builds on the data and infrastructure analyses, coherent improvement strategies, and the theory of action developed in Phase I. The plan developed in Phase II includes the activities, steps and resources required to implement the coherent improvement strategies, with attention to the research on implementation, timelines for implementation and measures needed to evaluate implementation and impact on the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families.

- Infrastructure Development: Specify improvements that will be made to the State infrastructure to better support EIS programs and/or EIS providers to implement and scale up evidence-based practices to improve the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families. Identify the steps the State will take to further align and leverage current improvement plans and initiatives in the State, including other early learning initiatives such as Race to the Top-Early Learning Challenge and the Home Visiting program, which impacts infants and toddlers with disabilities. This section must also identify who will be in charge of implementing the changes to infrastructure, resources needed, expected outcomes, and timelines for completing improvement efforts. In addition, the State should specify how it will involve multiple offices within the State lead agency (LA), as well as other State agencies (such as the State educational agency or SEA if different from the LA), in the improvement of its infrastructure.

- Support for EIS Program and/or EIS Provider Implementation of Evidence-Based Practices: Specify how the State will support EIS programs and/or EIS providers in implementing the evidence-based practices that will result in changes in LA, EIS program and/or EIS provider practices to achieve the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families. This section must identify steps and specific activities needed to implement the coherent improvement strategies, including communication strategies and stakeholder involvement; how identified barriers will be addressed; who will be in charge of implementing; how the activities will be implemented with fidelity; the resources that will be used to implement them; how the expected outcomes of the improvement strategies will be measured; and timelines for completion. In addition, the State should specify how it will involve multiple offices within the LA (or other State agencies including the SEA) to support EIS programs and/or EIS providers in scaling up and sustaining the implementation of the evidence-based practices once they have been implemented with fidelity.

- Evaluation: The evaluation must include short-term and long-term objectives to measure implementation of the SSIP and its impact on achieving measurable improvement in the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families and long-term objectives as those children exit Part C. The evaluation must be aligned to the theory of action and other components of the SSIP, include how stakeholders will be involved, and include the methods that the State will use to collect and analyze data to evaluate implementation and outcomes of the SSIP. The evaluation must specify how the State will use the information from the evaluation to examine the effectiveness of the implementation of the SSIP and the progress toward achieving intended improvements in the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families, and to make modifications to the SSIP as necessary, and how the information from the evaluation will be disseminated to stakeholders.

Phase III: Implementation and Evaluation

In Phase III, the State must, consistent with the evaluation described in Phase II, assess and report on its progress in implementing the SSIP. This will include data and analysis on the extent to which the State has made progress toward and/or met the State-established short-term and long-term objectives for implementation of the SSIP and its progress in achieving the State-identified Measurable Result(s) for Infants and Toddlers with Disabilities and their Families. If the State intends to continue implementing the SSIP without modifications, the State must describe how the data from the evaluation support this decision. Also, the State must provide a rationale for any revisions that have been made, or revisions the State plans to

make, in the SSIP in response to evaluation data, and describe how stakeholders were included in the decision-making process.

PAPERWORK BURDEN STATEMENT

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. It is estimated that each respondent will spend approximately 1,100 hours completing the APR. These estimates include time for reviewing instructions, searching any existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. The obligation to respond to this collection required to obtain or retain benefits (20 U.S.C. 1416(b)(1); 20 U.S.C. 1442; 20 U.S.C. 1416(b)(2)(C)(ii)). Send comments regarding the burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to the U.S. Department of Education, 400 Maryland Ave., SW, Washington, DC 20210-4537 or email ICDocketMgr@ed.gov and reference the OMB Control Number 1820-0578. Note: Please do not return the completed Part C SPP or APR forms to this address.

UNITED STATES DEPARTMENT OF EDUCATION

OFFICE OF SPECIAL EDUCATION AND
REHABILITATIVE SERVICES

DEC 2 3 2014

Contact Person
Name: Gregg Corr
Telephone: 202-245-7309

OSEP 15-06

MEMORANDUM

TO: Chief State School Officers, State Directors of Special Education, and State Data Managers

FROM: Melody Musgrove, Ed.D.
Director
Office of Special Education Programs (OSEP)

SUBJECT: Part B State Performance Plan (Part B - SPP) and Part B Annual Performance Report (Part B - APR)

**ACTION
REQUIRED:** Submission of the Part B - SPP and the Part B - APR by February 2, 2015

In accordance with 20 U.S.C. 1416(b)(1)(C) of the Individuals with Disabilities Education Act (IDEA or Part B), each State must amend, at least once every six years, its Part B State Performance Plan (Part B – SPP), and under 20 U.S.C. 1416(b)(2)(C)(ii)(II), report annually, through the Part B Annual Performance Report (Part B – APR) to the Secretary on the State's performance under its Part B – SPP. The Part B – SPP evaluates the State's efforts to implement the requirements and purposes of Part B and describes how the State will improve such implementation. The original Part B – SPP that all States submitted in 2005 covered a period of six years for Federal fiscal years (FFYs) 2005 through 2010 and OSEP extended the Part B – SPP for two years to cover FFYs 2011 and 2012.

On February 2, 2015, each State must submit a new Part B – SPP that covers the six year period for FFYs 2013 through 2018 and includes a new Indicator 17, the State Systemic Improvement Plan (SSIP) that is part of OSEP's Results Drive Accountability (RDA) Framework. In accordance with 20 U.S.C. 1416(b)(2)(C)(ii)(I) and 34 CFR §300.602(b)(1)(i)(A), each State must also report annually to the public on the performance of each local educational agency (LEA) located in the State on the targets in its Part B – SPP as soon as practicable, but no later than 120 days following the State's submission of its Part B – APR to the Secretary.

C-20

Information Collection 1820-0624, *Part B State Performance Plan (SPP)/Annual Performance
Report (APR)*, is available electronically at https://osep.grads360.org/#program/spp-apr-
resources and contains both Part B – SPP and Part B – APR instructions, and the document is
combined as one report and referred to throughout this memo as the SPP/APR. The link also
includes the *Part B Related Requirements* document as a reference. Beginning with the FFY
2013 SPP/APR, the State must submit its SPP/APR online using the SPP/APR module on
GRADS 360° (https://osep.grads360.org). The online SPP/APR module is only accessible to
State-authorized users who have a unique log-in user name and password. User names and
passwords are available to all State-authorized users by emailing EDEN_ss@ed.gov. States are
encouraged to make note of the due dates and implement a development schedule accordingly.

By February 2, 2015 (unless specifically indicated otherwise below), States must submit:

1) One report that includes both SPP and APR information.

2) An introduction, with sufficient detail to ensure that the Secretary and the public are
   informed of and understand the State's systems designed to drive improved results for
   children with disabilities and to ensure that the State educational agency (SEA) and local
   educational agencies (LEAs) meet the requirements of IDEA Part B. This introduction
   must include descriptions of the State's:

   a. General Supervision System: The systems that are in place to ensure that IDEA
      Part B requirements are met, e.g., monitoring, dispute resolution, etc.;

   b. Technical Assistance System: The mechanisms that the State has in place to
      ensure the timely delivery of high quality, evidenced based technical assistance
      and support to LEAs;

   c. Professional Development System: The mechanisms the State has in place to
      ensure that service providers have the skills to effectively provide services that
      improve results for children with disabilities;

   d. Stakeholder Involvement: The mechanism for soliciting broad stakeholder input
      on the State's targets in the SPP/APR and the development and implementation of
      new Indicator 17, the SSIP (Note that all of the information regarding Indicator
      17, including the baseline data for FFY 2013, targets for FFYs 2014 through
      2018, and other content is not due until April 1, 2015, as reflected in items 3, 4
      and 7 below) [1]; and

   e. Reporting to the Public: How and where the State reported to the public on the
      FFY 2012 performance of each LEA located in the State on the targets in the
      SPP/APR as soon as practicable, but no later than 120 days following the State's
      submission of its FFY 2012 APR, as required by 34 CFR §300.602(b)(1)(i)(A);
      and a description of where, on its Web site, a complete copy of the State's SPP,
      including any revision if the State has revised the SPP that it submitted with its
      FFY 2012 APR in 2014, is available.

3) Baseline data for FFY 2013 for Indicator 17.[2]

---

[1] As noted in the conference report to HR 1350, it is Congress' expectation that targets will be developed with broad
stakeholder input and will be disseminated to the public.
[2] For the FFY 2013 SPP/APR, all information related to Indicator 17, the SSIP, including baseline and targets, and
specific content, is due on April 1, 2015.

4) Annual State targets for Indicators 1 through 16, determined with stakeholder input, for each year and that cover the years of the SPP (i.e., FFY 2013 through FFY 2018). Targets for Indicator 17, which cover FFY 2014 through FFY 2018.

5) Data from FFY 2013 and other responsive APR information for Indicators 1 through 16.[3]

6) An explanation of any slippage in indicators where the State did not meet its FFY 2013 target.

7) The specific content required to complete Phase I of the SSIP required by Indicator 17. Phase I includes a detailed analysis that will guide the selection of coherent improvement strategies to increase State capacity to lead meaningful change in LEAs related to the State-identified Measurable Result(s) for Children with Disabilities. (See Indicator 17 for specific content of Phases II and III of the SSIP, which must be included with the States' FFY 2014 through FFY 2018 SPP/APRs.):

   a. Data Analysis

   b. Analysis of State Infrastructure to Support Improvement and Build Capacity

   c. State-identified Measurable Result(s) for Children with Disabilities

   d. Selection of Coherent Improvement Strategies

   e. Theory of Action

8) Information to address any required actions identified in OSEP's letter responding to the State's February 3, 2014 submission of its FFY 2012 APR, including providing information in an attachment to address required actions identified in previous Indicator 15 (Timely Correction).

When completing the SPP/APR, States will use the following parts of Information Collection 1820-0624:

- *SPP/APR Instructions*

- *Part B Indicator Measurement Table* with Instructions

The *Part B Indicator Measurement Table* lists the Monitoring Priorities and Indicators, required data sources and measurement and instructions for providing the required information for each indicator. In addition to the percentages required in the indicators, States are required to provide <u>actual numbers</u> used in the calculations.[4]

The Department will review the information provided in the State's FFY 2013 APR, other State-reported data, information obtained through monitoring visits, and other public information, and will determine, under IDEA section 616(d), if the State meets the requirements and purposes of Part B of the IDEA or needs assistance, needs intervention, or needs substantial intervention in implementing the requirements of Part B of the IDEA. When reporting on correction of noncompliance in the APR, each State must include confirmation that: (1) each LEA has corrected each individual instance of child-specific noncompliance; and (2) the State has verified that the LEA is currently correctly implementing the specific regulatory requirement. The State should review IDEA section 616(e) regarding the potential enforcement actions the Department

---

[3] Previous Indicator 15 (Timely Correction) and Indicator 20 (Timely and Accurate Data) have been removed. Consequently, the indicators have been renumbered.

[4] SEAs are not required to provide the actual numbers for Indicators 5, 6, 15, and 16 because they are using 618 State-reported data for these indicators.

is required to take as a result of, and potential future impact of, the Department's annual determination.

Prior to issuing the Department's 2015 determination for each State, OSEP will offer the State the opportunity to clarify or correct the data submitted in its FFY 2013 SPP/APR. OSEP will provide feedback through the online SPP/APR module on the State's February 2, 2015 FFY 2013 APR data that OSEP will consider in making the Department's determination under IDEA section 616(d). In response to OSEP's feedback, the State must submit to OSEP through the online SPP/APR module any corrected data and clarify any misunderstandings by OSEP about the data submitted. In order for the State's clarifications or corrections to data to be considered, the State's response must be certified as completed no later than close-of-business on the date specified in OSEP's online SPP/APR module summary.

Please note that any State that does not meet the February 2, 2015[5] timeline for submitting data for each indicator will not be permitted to submit clarifications or corrections for purposes of the Department's determination under IDEA section 616(d).

In accordance with section 616(e) of the IDEA and 34 CFR §300.604, in the Department's June 23, 2014 determination letters, the Secretary advised States that were needs assistance for two or more consecutive years of available sources of technical assistance. The Secretary directed such a State to determine the results element(s) and/or compliance indicator(s), and improvement strategies, on which it would focus its use of available technical assistance in order to improve its performance. The Secretary strongly encouraged each State to access technical assistance related to those results element(s) and compliance indicator(s) for which the State received a score of zero on the 2014 Compliance Matrix. These States must report with their FFY 2013 APR submission (due February 2, 2015) on: (1) the technical assistance sources from which the State received assistance; and (2) what actions the State took as a result of that technical assistance. The extent to which your State takes advantage of available technical assistance may affect the actions we take under section 616, should your State not be identified as meets requirements in 2015.

If you have any further questions about the Part B – SPP and/or Part B – APR and/or the submission options listed above, please contact your Part B State Contact.

cc:　National Center for Systemic Improvement (NCSI)
　　　Early Childhood Technical Assistance (ECTA) Center

---

[5] See Note 2

Part B State Performance Plan (SPP) and Annual Performance Report (APR)

Part B Indicator Measurement Table

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
Monitoring Priority: FAPE in the LRE		
1. Percent of youth with IEPs graduating from high school with a regular diploma. (20 U.S.C. 1416 (a)(3)(A))	**Data Source:** Same data as used for reporting to the Department under Title I of the Elementary and Secondary Education Act (ESEA). **Measurement:** States must report using the adjusted cohort graduation rate required under the ESEA.	*Sampling is not allowed.* Describe the results of the State's examination of the data for the year before the reporting year (e.g., for the FFY 2013 APR, use data from 2012-2013), and compare the results to the target. Provide the actual numbers used in the calculation. Provide a narrative that describes the conditions youth must meet in order to graduate with a regular diploma and, if different, the conditions that youth with IEPs must meet in order to graduate with a regular diploma. If there is a difference, explain why. Targets should be the same as the annual graduation rate targets under Title I of the ESEA.
2. Percent of youth with IEPs dropping out of high school. (20 U.S.C. 1416 (a)(3)(A))	**OPTION 1:** **Data Source:** Same data as used for reporting to the Department under IDEA section 618. **Measurement:** States must report a percentage using the number of youth with IEPs (ages 14-21) who exited special education due to dropping out in the numerator and the number of all youth with IEPs who left high school (ages 14-21) in the denominator. **OPTION 2:** Use same data source and measurement that the State used to report in its FFY 2010 APR that was	*Sampling is not allowed.* **OPTION 1:** Use 618 exiting data reported to the Department, and the definitions specified, in the EDFacts file specification C009. **http://www2.ed.gov/about/inits/ed/edfacts/eden/non-xml/c009-9-0.doc** Include in the denominator the following exiting categories: (a) graduated with a regular high school diploma; (b) received a certificate; (c) reached maximum age; (d) dropped out; or (e) died. Do not include in the denominator the number of youths with IEPs who exited special education due to: (a) transferring to regular education; or (b) who

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
	submitted on February 1, 2012.	moved, but are known to be continuing in education. **OPTION 2:** Use the annual event school dropout rate for students leaving a school in a single year determined in accordance with the National Center for Education Statistic's Common Core of Data. Data for this indicator are "lag" data. Describe the results of the State's examination of the data for the year before the reporting year (e.g., for the FFY 2013 APR, use data from 2012-2013), and compare the results to the target. Provide a narrative that describes what counts as dropping out for all youth and, if different, what counts as dropping out for youth with IEPs. If there is a difference, explain why.

3. Participation and performance of children with IEPs on statewide assessments:
 	
A. Percent of the districts with a disability subgroup that meets the State's minimum "n" size that meet the State's AYP/AMO targets for the disability subgroup.
 	
B. Participation rate for children with IEPs.
 	
C. Proficiency rate for children with IEPs against grade level, modified and alternate academic achievement standards.
 	
(20 U.S.C. 1416 (a)(3)(A))
	Data Source:
	
	3A. (choose either 3A.1 or 3A.2)
	
	  3A.1 AYP data used for accountability reporting under Title I of the ESEA.
	
	  3A.2 AMO data used for accountability reporting under Title I of the ESEA as a result of ESEA flexibility.
	
	3B. Assessment data reported in the Consolidated State Performance Report (CSPR) reporting on ESEA (EDFacts file specifications C185 and 188).
	
	3C. Assessment data reported in the Consolidated State Performance Report (CSPR) reporting on ESEA (EDFacts file specifications C175 and 178).
	
	Measurement:
	
	A. (choose either A.1 or A.2)
	
	  A.1 AYP percent = [(# of districts with a disability subgroup that meets the State's minimum "n" size
		Describe the results of the calculations and compare the results to the targets. Provide the actual numbers used in the calculation.
		
		States are encouraged to present their APR information in summary tables and include multiple years of data for comparison purposes.
		
		Include information regarding where to find public reports of assessment results, i.e., link to the Web site where results are reported.
		
		Indicator 3A: The data source and measurement for 3A is dependent on whether the State applied for, and was granted, a waiver of the requirements to determine Adequate Yearly Progress (AYP) for LEAs and schools as part of requesting ESEA flexibility. States that either did not apply for and receive ESEA flexibility, or applied for and received that flexibility but did not apply for a waiver of determining AYP should choose data source and measurement 3A.1. States with an approved ESEA flexibility request that includes a waiver of determining AYP should choose

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
	that meet the State's AYP targets for the disability subgroup) divided by the (total # of districts that have a disability subgroup that meets the State's minimum "n" size)] times 100. A.2 AMO percent = [(# of districts with a disability subgroup that meets the State's minimum "n" size that meet the State's AMO targets for the disability subgroup) divided by the (total # of districts that have a disability subgroup that meets the State's minimum "n" size)] times 100. B. Participation rate percent = [(# of children with IEPs participating in an assessment) divided by the (total # of children with IEPs enrolled during the testing window, calculated separately for reading and math)]. The participation rate is based on all children with IEPs, including both children with IEPs enrolled for a full academic year and those not enrolled for a full academic year. C. Proficiency rate percent = ([(# of children with IEPs scoring at or above proficient against grade level, modified and alternate academic achievement standards) divided by the (total # of children with IEPs who received a valid score and for whom a proficiency level was assigned, and, calculated separately for reading and math)]. The proficiency rate includes both children with IEPs enrolled for a full academic year and those not enrolled for a full academic year.	data source and measurement 3A.2. Report only on the AYP/AMO assessment targets for reading/language arts and mathematics proficiency, not targets for graduation or other elements of AYP/AMO. Indicator 3B: Provide separate reading/language arts and mathematics participation rates, inclusive of all ESEA grades assessed (3-8 and high school), for children with IEPs. Account for ALL children with IEPs, in all grades assessed, including children not participating in assessments and those not enrolled for a full academic year. Only include children with disabilities who had an IEP at the time of testing. Indicator 3C: Proficiency calculations in this APR must result in proficiency rates for each content area across all ESEA assessments (combining regular and all alternates) for children with IEPs, in all grades assessed (3-8 and high school), including both children with IEPs enrolled for a full academic year and those not enrolled for a full academic year. States are encouraged to report using two rates – one for reading/language arts covering all assessed grades and one for mathematics covering all assessed grades. Only include children with disabilities who had an IEP at the time of testing.
4. Rates of suspension and expulsion: A. Percent of districts that have a significant discrepancy in the rate of suspensions and expulsions of greater than 10 days in a school year for children with IEPs; and B. Percent of districts that have: (a) a significant discrepancy, by race or	**Data Source:** Data collected under section 618 of the Individuals with Disabilities Education Act (IDEA) (Report of Children with Disabilities Subject to Disciplinary Removal). Discrepancy can be computed by either comparing the rates of suspensions and expulsions for children with IEPs to rates for nondisabled children within the LEA or by comparing the rates of	*Sampling from State's 618 data is not allowed.* Describe the results of the State's examination of the data for the year before the reporting year (e.g., for the FFY 2013 APR, use data from 2012-2013), including data disaggregated by race and ethnicity to determine if significant discrepancies are occurring in the rates of long-term suspensions and expulsions of children with IEPs, as required at 20 U.S.C.

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
ethnicity, in the rate of suspensions and expulsions of greater than 10 days in a school year for children with IEPs; and (b) policies, procedures or practices that contribute to the significant discrepancy and do not comply with requirements relating to the development and implementation of IEPs, the use of positive behavioral interventions and supports, and procedural safeguards. (20 U.S.C. 1416(a)(3)(A); 1412(a)(22))	suspensions and expulsions for children with IEPs among LEAs within the State. **Measurement:** A. Percent = [(# of districts that have a significant discrepancy in the rates of suspensions and expulsions for greater than 10 days in a school year of children with IEPs) divided by the (# of districts in the State)] times 100. B. Percent = [(# of districts that have: (a) a significant discrepancy, by race or ethnicity, in the rates of suspensions and expulsions of greater than 10 days in a school year of children with IEPs; and (b) policies, procedures or practices that contribute to the significant discrepancy and do not comply with requirements relating to the development and implementation of IEPs, the use of positive behavioral interventions and supports, and procedural safeguards) divided by the (# of districts in the State)] times 100. Include State's definition of "significant discrepancy."	1412(a)(22). The State's examination must include one of the following comparisons: - The rates of suspensions and expulsions for children with IEPs among LEAs within the State; or - The rates of suspensions and expulsions for children with IEPs to nondisabled children within the LEAs. In the description, specify which method the State used to determine possible discrepancies and explain what constitutes those discrepancies. If the State used a minimum "n" size requirement, report the number of districts excluded from the calculation as a result of this requirement. States have the option of using the "total number of districts" OR the "number of districts that meet the State's minimum n size for one or more racial/ethnic group" as the denominator in the calculation for B4A or B4B. For 4A, provide the actual numbers used in the calculation and if significant discrepancies occurred, describe how the State educational agency reviewed and, if appropriate, revised (or required the affected local educational agency to revise) its policies, procedures, and practices relating to the development and implementation of IEPs, the use of positive behavioral interventions and supports, and procedural safeguards, to ensure that such policies, procedures, and practices comply with applicable requirements. For 4B, provide the following: (a) the number of districts that have a significant discrepancy, by race or ethnicity, in the rates of suspensions and expulsions of greater than 10 days in a school year for children with IEPs; and (b) the number of districts in which policies, procedures or practices contribute to the significant discrepancy and do not comply with requirements relating to the development and

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
		implementation of IEPs, the use of positive behavioral interventions and supports, and procedural safeguards. If discrepancies occurred and the district with discrepancies had policies, procedures or practices that contributed to the significant discrepancy and that do not comply with requirements relating to the development and implementation of IEPs, the use of positive behavioral interventions and supports, and procedural safeguards, describe how the State ensured that such policies, procedures, and practices were revised to comply with applicable requirements consistent with OSEP Memorandum 09-02, dated October 17, 2008. Targets must be 0% for 4B.
5. Percent of children with IEPs aged 6 through 21 served: A. Inside the regular class 80% or more of the day; B. Inside the regular class less than 40% of the day; and C. In separate schools, residential facilities, or homebound/hospital placements. (20 U.S.C. 1416(a)(3)(A))	**Data Source:** Data collected under IDEA section 618. **Measurement:** A. Percent = [(# of children with IEPs served inside the regular class 80% or more of the day) divided by the (total # of students aged 6 through 21 with IEPs)] times 100. B. Percent = [(# of children with IEPs served inside the regular class less than 40% of the day) divided by the (total # of students aged 6 through 21 with IEPs)] times 100. C. Percent = [(# of children with IEPs served in separate schools, residential facilities, or homebound/hospital placements) divided by the (total # of students aged 6 through 21 with IEPs)] times 100.	*For this indicator, report 618 data that were collected on a date between October 1 and December 1, 2013 and due on February 1, 2014. Sampling from State's 618 data is not allowed.* Describe the results of the calculations and compare the results to the target. If the data reported in this indicator are not the same as the State's data reported under IDEA section 618, explain.
6. Percent of children aged 3 through 5 with IEPs attending a:	**Data Source:** Data collected under IDEA section 618.	*For this indicator, report 618 data that were collected on a date between October 1 and December 1, 2013 and due on February 1, 2014. Sampling from*

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
A. Regular early childhood program and receiving the majority of special education and related services in the regular early childhood program; and B. Separate special education class, separate school or residential facility. (20 U.S.C. 1416(a)(3)(A))	**Measurement:** A. Percent = [(# of children aged 3 through 5 with IEPs attending a regular early childhood program and receiving the majority of special education and related services in the regular early childhood program) divided by the (total # of children aged 3 through 5 with IEPs)] times 100. B. Percent = [(# of children aged 3 through 5 with IEPs attending a separate special education class, separate school or residential facility) divided by the (total # of children aged 3 through 5 with IEPs)] times 100.	*State's 618 data is not allowed.* If the data reported in this indicator are not the same as the State's data reported under IDEA section 618, explain.
7. Percent of preschool children aged 3 through 5 with IEPs who demonstrate improved: A. Positive social-emotional skills (including social relationships); B. Acquisition and use of knowledge and skills (including early language/ communication and early literacy); and C. Use of appropriate behaviors to meet their needs. (20 U.S.C. 1416 (a)(3)(A))	**Data Source:** State selected data source. **Measurement:** Outcomes: A. Positive social-emotional skills (including social relationships); B. Acquisition and use of knowledge and skills (including early language/communication and early literacy); and C. Use of appropriate behaviors to meet their needs. Progress categories for A, B and C: a. Percent of preschool children who did not improve functioning = [(# of preschool children who did not improve functioning) divided by (# of preschool children with IEPs assessed)] times 100. b. Percent of preschool children who improved functioning but not sufficient to move nearer to functioning comparable to same-aged peers = [(# of preschool children who	*Sampling of **children for assessment** is allowed. When sampling is used, submit a description of the sampling methodology outlining how the design will yield valid and reliable estimates. (See General Instructions on page 2 for additional instructions on sampling.)* Describe the results of the calculations and compare the results to the targets. States will use the progress categories for each of the three Outcomes to calculate and report the two Summary Statements. States have provided targets for the two Summary Statements for the three Outcomes (six numbers for targets for each FFY). Report progress data and calculate Summary Statements to compare against the six targets. Provide the actual numbers and percentages for the five reporting categories for each of the three outcomes. In presenting results, provide the criteria for defining "comparable to same-aged peers." If a State is using the Early Childhood Outcomes Center (ECO) Child Outcomes Summary Form (COSF), then the criteria for defining "comparable to same-aged peers" has been defined as a child who has been assigned a

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
	improved functioning but not sufficient to move nearer to functioning comparable to same-aged peers) divided by (# of preschool children with IEPs assessed)] times 100. c. Percent of preschool children who improved functioning to a level nearer to same-aged peers but did not reach it = [(# of preschool children who improved functioning to a level nearer to same-aged peers but did not reach it) divided by (# of preschool children with IEPs assessed)] times 100. d. Percent of preschool children who improved functioning to reach a level comparable to same-aged peers = [(# of preschool children who improved functioning to reach a level comparable to same-aged peers) divided by (# of preschool children with IEPs assessed)] times 100. e. Percent of preschool children who maintained functioning at a level comparable to same-aged peers = [(# of preschool children who maintained functioning at a level comparable to same-aged peers) divided by (# of preschool children with IEPs assessed)] times 100. **Summary Statements for Each of the Three Outcomes:** **Summary Statement 1:** Of those preschool children who entered or exited the preschool program below age expectations in each Outcome, the percent who substantially increased their rate of growth by the time they turned 6 years of age or exited the program. **Measurement for Summary Statement 1:** Percent = # of preschool children reported in progress category (c) plus # of preschool children	scored of 6 or 7 on the COSF. In addition, list the instruments and procedures used to gather data for this indicator, including if the State is using the ECO COSF.

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
	reported in category (d) divided by [# of preschool children reported in progress category (a) plus # of preschool children reported in progress category (b) plus # of preschool children reported in progress category (c) plus # of preschool children reported in progress category (d)] times 100. **Summary Statement 2:** The percent of preschool children who were functioning within age expectations in each Outcome by the time they turned 6 years of age or exited the program. **Measurement for Summary Statement 2:** Percent = # of preschool children reported in progress category (d) plus # of preschool children reported in progress category (e) divided by [the total # of preschool children reported in progress categories (a) + (b) + (c) + (d) + (e)] times 100.	
8. Percent of parents with a child receiving special education services who report that schools facilitated parent involvement as a means of improving services and results for children with disabilities. (20 U.S.C. 1416(a)(3)(A))	**Data Source:** State selected data source. **Measurement:** Percent = [(# of respondent parents who report schools facilitated parent involvement as a means of improving services and results for children with disabilities) divided by the (total # of respondent parents of children with disabilities)] times 100.	*Sampling **of parents from whom response is requested** is allowed. When sampling is used, submit a description of the sampling methodology outlining how the design will yield valid and reliable estimates. (See General Instructions on page 2 for additional instructions on sampling.)* Describe the results of the calculations and compare the results to the target. Include a description of how the State has ensured that the response data are valid and reliable, including how the data represent the demographics of the State. Provide the actual numbers used in the calculation. If the State is using a separate data collection methodology for preschool children, the State must provide separate baseline data, targets, and actual target data or discuss the procedures used to combine data from school age and preschool data collection methodologies in a manner that is valid and reliable. While a survey is not required for this indicator, a

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
		State using a survey must submit a copy of any new or revised survey with its APR. States are encouraged to work in collaboration with their OSEP-funded parent centers in collecting data.
Monitoring Priority: Disproportionate Representation		
9. Percent of districts with disproportionate representation of racial and ethnic groups in special education and related services that is the result of inappropriate identification. (20 U.S.C. 1416(a)(3)(C))	**Data Source:** Data collected under IDEA section 618 (Report of Children with Disabilities Receiving Special Education Under Part B of the Individuals with Disabilities Education Act, As Amended) and the State's analysis to determine if the disproportionate representation of racial and ethnic groups in special education and related services was the result of inappropriate identification. **Measurement:** Percent = [(# of districts with disproportionate representation of racial and ethnic groups in special education and related services that is the result of inappropriate identification) divided by the (# of districts in the State)] times 100. Include State's definition of "disproportionate representation." Based on its review of the 618 data for FFY 2013, describe how the State made its annual determination that the disproportionate overrepresentation it identified of racial and ethnic groups in special education and related services was the result of inappropriate identification as required by §§300.600(d)(3) and 300.602(a), e.g., using monitoring data; reviewing policies, practices and procedures, etc. In determining disproportionate representation, analyze data, for each district, for all racial and ethnic groups in the district, or all racial and ethnic groups in the district that meet a minimum	Provide racial/ethnic disproportionality data for children aged 6 through 21 served under IDEA. Provide these data for all children with disabilities. States are not required to report on underrepresentation. Provide the number of districts identified with disproportionate representation of racial and ethnic groups in special education and related services and the number of districts identified with disproportionate representation that is the result of inappropriate identification. Consider using multiple methods in calculating disproportionate representation of racial and ethnic groups to reduce the risk of overlooking potential problems. Describe the method(s) used to calculate disproportionate representation. If the State used a minimum "n" size requirement, report the number of districts totally excluded from the calculation as a result of this requirement because the district did not meet the minimum "n" size for any racial/ethnic group. States have the option of using the "total number of districts" OR the "number of districts that meet the State's minimum "n" size for one or more racial/ethnic group" as the denominator in the calculation. Targets must be 0%. Provide detailed information about the timely correction of noncompliance as noted in OSEP's

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
	'n' size set by the State. Report on the percent of districts in which disproportionate representation of racial and ethnic groups in special education and related services is the result of inappropriate identification, even if the determination of inappropriate identification was made after the end of the FFY 2013 reporting period, i.e., after June 30, 2014. If inappropriate identification is identified, report on corrective actions taken.	response table for the previous APR. If the State did not ensure timely correction of the previous noncompliance, provide information on the extent to which noncompliance was subsequently corrected (more than one year after identification). In addition, provide information regarding the nature of any continuing noncompliance, improvement activities completed (e.g., review of policies and procedures, technical assistance, training, etc.) and any enforcement actions that were taken.
10. Percent of districts with disproportionate representation of racial and ethnic groups in specific disability categories that is the result of inappropriate identification. (20 U.S.C. 1416(a)(3)(C))	**Data Source:** Data collected under IDEA section 618 (Report of Children with Disabilities Receiving Special Education Under Part B of the Individuals with Disabilities Education Act, As Amended) and the State's analysis to determine if the disproportionate representation of racial and ethnic groups in specific disability categories was the result of inappropriate identification. **Measurement:** Percent = [(# of districts with disproportionate representation of racial and ethnic groups in specific disability categories that is the result of inappropriate identification) divided by the (# of districts in the State)] times 100. Include State's definition of "disproportionate representation." Based on its review of the 618 data for FFY 2013, describe how the State made its annual determination that the disproportionate overrepresentation it identified of racial and ethnic groups in specific disability categories was the result of inappropriate identification as required by §§300.600(d)(3) and 300.602(a), e.g., using monitoring data; reviewing policies, practices and	Provide racial/ethnic disproportionality data for children aged 6 through 21 served under IDEA. Provide these data at a minimum for children in the following six disability categories: intellectual disability, specific learning disabilities, emotional disturbance, speech or language impairments, other health impairments, and autism. If a State has identified disproportionate representation of racial and ethnic groups in specific disability categories other than these six disability categories, the State must include these data and report on whether the State determined that the disproportionate representation of racial and ethnic groups in specific disability categories was the result of inappropriate identification. States are not required to report on underrepresentation. Provide the number of districts identified with disproportionate representation of racial and ethnic groups in specific disability categories and the number of districts identified with disproportionate representation that is the result of inappropriate identification. If the State used a minimum "n" size requirement, report the number of districts totally excluded from the calculation as a result of this requirement because the district did not meet the minimum "n" size for any racial/ethnic group. States

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
	procedures, etc. In determining disproportionate representation, analyze data, for each district, for all racial and ethnic groups in the district, or all racial and ethnic groups in the district that meet a minimum 'n' size set by the State. Report on the percent of districts in which disproportionate representation of racial and ethnic groups in specific disability categories is the result of inappropriate identification, even if the determination of inappropriate identification was made after the end of the FFY 2013, i.e., after June 30, 2014. If inappropriate identification is identified, report on corrective actions taken.	have the option of using the "total number of districts" OR the "number of districts that meet the State's minimum n size for one or more racial/ethnic group" as the denominator in the calculation. Consider using multiple methods in calculating disproportionate representation of racial and ethnic groups to reduce the risk of overlooking potential problems. Describe the method(s) used to calculate disproportionate representation. Targets must be 0%. Provide detailed information about the timely correction of noncompliance as noted in OSEP's response table for the previous APR. If the State did not ensure timely correction of the previous noncompliance, provide information on the extent to which noncompliance was subsequently corrected (more than one year after identification). In addition, provide information regarding the nature of any continuing noncompliance, improvement activities completed (e.g., review of policies and procedures, technical assistance, training, etc.) and any enforcement actions that were taken.

Monitoring Priority: Effective General Supervision Part B

Effective General Supervision Part B / Child Find

11. Percent of children who were evaluated within 60 days of receiving parental consent for initial evaluation or, if the State establishes a timeframe within which the evaluation must be conducted, within that timeframe. (20 U.S.C. 1416(a)(3)(B))	**Data Source:** Data to be taken from State monitoring or State data system and must be based on actual, not an average, number of days. Indicate if the State has established a timeline and, if so, what is the State's timeline for initial evaluations. **Measurement:** a. # of children for whom parental consent to evaluate was received.	*If data are from State monitoring, describe the method used to select LEAs for monitoring. If data are from a State database, include data for the entire reporting year.* Describe the results of the calculations and compare the results to the target. Describe the method used to collect these data, and if data are from the State's monitoring, describe the procedures used to collect these data. Provide the actual numbers used in the calculation.

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
	b. # of children whose evaluations were completed within 60 days (or State-established timeline). Account for children included in (a), but not included in (b). Indicate the range of days beyond the timeline when the evaluation was completed and any reasons for the delays. Percent = [(b) divided by (a)] times 100.	Note that under 34 CFR §300.301(d), the timeframe set for initial evaluation does not apply to a public agency if: (1) The parent of a child repeatedly fails or refuses to produce the child for the evaluation; or (2) A child enrolls in a school of another public agency after the timeframe for initial evaluations has begun, and prior to a determination by the child's previous public agency as to whether the child is a child with a disability. States should not report these exceptions in either the numerator (b) or denominator (a). If the State-established timeframe provides for exceptions through State regulation or policy, describe cases falling within those exceptions and include in b. Targets must be 100%. Provide detailed information about the timely correction of noncompliance as noted in OSEP's response table for the previous APR. If the State did not ensure timely correction of the previous noncompliance, provide information on the extent to which noncompliance was subsequently corrected (more than one year after identification). In addition, provide information regarding the nature of any continuing noncompliance, improvement activities completed (e.g., review of policies and procedures, technical assistance, training, etc.) and any enforcement actions that were taken.
Effective General Supervision Part B / Effective Transition		
12. Percent of children referred by Part C prior to age 3, who are found eligible for Part B, and who have an IEP developed and implemented by their third birthdays. (20 U.S.C. 1416(a)(3)(B))	**Data Source:** Data to be taken from State monitoring or State data system. **Measurement:** a. # of children who have been served in Part C and referred to Part B for Part B eligibility determination. b. # of those referred determined to be NOT eligible	*If data are from State monitoring, describe the method used to select LEAs for monitoring. If data are from a State database, include data for the entire reporting year.* Describe the results of the calculations and compare the results to the target. Describe the method used to collect these data and if data are from the State's monitoring, describe the procedures used to collect these data. Provide the actual numbers used in the

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
	and whose eligibility was determined prior to their third birthdays. c. # of those found eligible who have an IEP developed and implemented by their third birthdays. d. # of children for whom parent refusal to provide consent caused delays in evaluation or initial services or to whom exceptions under 34 CFR §300.301(d) applied. e. # of children determined to be eligible for early intervention services under Part C less than 90 days before their third birthdays. Account for children included in (a), but not included in b, c, d, or e. Indicate the range of days beyond the third birthday when eligibility was determined and the IEP developed, and the reasons for the delays. Percent = [(c) divided by (a - b - d - e)] times 100.	calculation. Targets must be 100%. Provide detailed information about the timely correction of noncompliance as noted in OSEP's response table for the previous APR. If the State did not ensure timely correction of the previous noncompliance, provide information on the extent to which noncompliance was subsequently corrected (more than one year after identification). In addition, provide information regarding the nature of any continuing noncompliance, improvement activities completed (e.g., review of policies and procedures, technical assistance, training, etc.) and any enforcement actions that were taken.
13. Percent of youth with IEPs aged 16 and above with an IEP that includes appropriate measurable postsecondary goals that are annually updated and based upon an age appropriate transition assessment, transition services, including courses of study, that will reasonably enable the student to meet those postsecondary goals, and annual IEP goals related to the student's transition services needs. There also must be evidence that the student was invited to the IEP Team meeting where transition services are to be discussed and evidence that, if appropriate, a representative of any participating agency was invited to the IEP Team meeting with the prior consent of the parent or student who has reached the age of majority.	**Data Source:** Data to be taken from State monitoring or State data system. **Measurement:** Percent = [(# of youth with IEPs aged 16 and above with an IEP that includes appropriate measurable postsecondary goals that are annually updated and based upon an age appropriate transition assessment, transition services, including courses of study, that will reasonably enable the student to meet those postsecondary goals, and annual IEP goals related to the student's transition services needs. There also must be evidence that the student was invited to the IEP Team meeting where transition services are to be discussed and evidence that, if appropriate, a representative of any participating agency was invited to the IEP Team meeting with the prior consent of the parent or student who has reached the age of majority) divided by the (# of	*If data are from State monitoring, describe the method used to select LEAs for monitoring. If data are from a State database, include data for the entire reporting year.* Describe the results of the calculations and compare the results to the target. Describe the method used to collect these data and if data are from the State's monitoring, describe the procedures used to collect these data. Provide the actual numbers used in the calculation. Targets must be 100%. Provide detailed information about the timely correction of noncompliance as noted in OSEP's response table for the previous APR. If the State did not ensure timely correction of the previous noncompliance, provide information on the extent to which noncompliance was subsequently corrected (more than one year after identification). In addition, provide information regarding the nature of any

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
(20 U.S.C. 1416(a)(3)(B))	youth with an IEP age 16 and above)] times 100.	continuing noncompliance, improvement activities completed (e.g., review of policies and procedures, technical assistance, training, etc.) and any enforcement actions that were taken.
14. Percent of youth who are no longer in secondary school, had IEPs in effect at the time they left school, and were: A. Enrolled in higher education within one year of leaving high school. B. Enrolled in higher education or competitively employed within one year of leaving high school. C. Enrolled in higher education or in some other postsecondary education or training program; or competitively employed or in some other employment within one year of leaving high school. (20 U.S.C. 1416(a)(3)(B))	**Data Source:** State selected data source. **Measurement:** A. Percent enrolled in higher education = [(# of youth who are no longer in secondary school, had IEPs in effect at the time they left school and were enrolled in higher education within one year of leaving high school) divided by the (# of respondent youth who are no longer in secondary school and had IEPs in effect at the time they left school)] times 100. B. Percent enrolled in higher education or competitively employed within one year of leaving high school = [(# of youth who are no longer in secondary school, had IEPs in effect at the time they left school and were enrolled in higher education or competitively employed within one year of leaving high school) divided by the (# of respondent youth who are no longer in secondary school and had IEPs in effect at the time they left school)] times 100. C. Percent enrolled in higher education, or in some other postsecondary education or training program; or competitively employed or in some other employment = [(# of youth who are no longer in secondary school, had IEPs in effect at the time they left school and were enrolled in higher education, or in some other postsecondary education or training program; or competitively employed or in some other employment) divided by the (# of respondent youth who are no longer in secondary school and had IEPs in effect at the time they left school)] times 100.	*Sampling **of youth who had IEPs and are no longer in secondary school** is allowed. When sampling is used, submit a description of the sampling methodology outlining how the design will yield valid and reliable estimates of the target population. (See* General Instructions *on page 2 for additional instructions on sampling.)* Collect data by September 2014 on students who left school during 2012-2013, timing the data collection so that at least one year has passed since the students left school. Include students who dropped out during 2012-2013 or who were expected to return but did not return for the current school year. This includes all youth who had an IEP in effect at the time they left school, including those who graduated with a regular diploma or some other credential, dropped out, or aged out. **I. *Definitions*** Enrolled in higher education as used in measures A, B, and C means youth have been enrolled on a full- or part-time basis in a community college (two year program) or college/university (four or more year program) for at least one complete term, at any time in the year since leaving high school. Competitive employment as used in measures B and C means that youth have worked for pay at or above the minimum wage in a setting with others who are nondisabled for a period of 20 hours a week for at least 90 days at any time in the year since leaving high school. This includes military employment. Enrolled in other postsecondary education or training as used in measure C, means youth have been

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
		enrolled on a full- or part-time basis for at least 1 complete term at any time in the year since leaving high school in an education or training program (e.g., Job Corps, adult education, workforce development program, vocational technical school which is less than a two year program).
		Some other employment as used in measure C means youth have worked for pay or been self-employed for a period of at least 90 days at any time in the year since leaving high school. This includes working in a family business (e.g., farm, store, fishing, ranching, catering services, etc.).
		II. Data Reporting
		Provide the actual numbers for each of the following mutually exclusive categories. The actual number of "leavers" who are:
		1. Enrolled in higher education within one year of leaving high school;
		2. Competitively employed within one year of leaving high school (but not enrolled in higher education);
		3. Enrolled in some other postsecondary education or training program within one year of leaving high school (but not enrolled in higher education or competitively employed);
		4. In some other employment within one year of leaving high school (but not enrolled in higher education, some other postsecondary education or training program, or competitively employed).
		"Leavers" should only be counted in one of the above categories, and the categories are organized hierarchically. So, for example, "leavers" who are enrolled in full- or part-time higher education within one year of leaving high school should only be reported in category 1, even if they also happen to be employed. Likewise, "leavers" who are not

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
		enrolled in either part- or full-time higher education, but who are competitively employed, should only be reported under category 2, even if they happen to be enrolled in some other postsecondary education or training program.
		III. Reporting On the Measures/Indicators
		Targets must be established for measures A, B, and C.
		Measure A: For purposes of reporting on the measures/indicators, please note that any youth enrolled in an institution of higher education (that meets any definition of this term in the Higher Education Act (HEA)) within one year of leaving high school *must* be reported under measure A. This could include youth who also happen to be competitively employed, or in some other training program; however, the key outcome we are interested in here is enrollment in higher education.
		Measure B: All youth reported under measure A should also be reported under measure B, in addition to all youth that obtain competitive employment within one year of leaving high school.
		Measure C: All youth reported under measures A and B should also be reported under measure C, in addition to youth that are enrolled in some other postsecondary education or training program, or in some other employment.
		Describe the calculations and results using actual numbers and compare these results to the targets. Include a description of how the State has ensured that survey data are valid and reliable, including how the data represent the demographics of the State.
Effective General Supervision Part B / General Supervision		
15. Percent of hearing requests that went to	**Data Source:**	*Sampling is not allowed.*

Monitoring Priorities and Indicators	Data Source and Measurement	Instructions for Indicators/Measurement
resolution sessions that were resolved through resolution session settlement agreements. (20 U.S.C. 1416(a)(3(B))	Data collected under IDEA section 618. **Measurement:** Percent = (3.1(a) divided by 3.1) times 100.	Describe the results of the calculations and compare the results to the target. States are not required to establish baseline or targets if the number of resolution sessions is less than 10. In a reporting period when the number of resolution sessions reaches 10 or greater, develop baseline, targets and improvement activities, and report on them in the corresponding APR. States may express their targets in a range, e.g., 75-85%. If the data reported in this indicator are not the same as the State's data under IDEA section 618, explain. States are not required to report data at the LEA level.
16. Percent of mediations held that resulted in mediation agreements. (20 U.S.C. 1416(a)(3(B))	**Data Source:** Data collected under IDEA section 618. **Measurement:** Percent = [(2.1(a)(i) + 2.1(b)(i)) divided by 2.1] times 100.	*Sampling is not allowed.* Describe the results of the calculations and compare the results to the target. States are not required to establish baseline or targets if the number of mediations is less than 10. In a reporting period when the number of mediations reaches ten or greater, develop baseline, targets and improvement activities, and report on them in the corresponding APR. States may express their targets in a range, e.g., 75-85%. If the data reported in this indicator are not the same as the State's data under IDEA section 618, explain. States are not required to report data at the LEA level.

MONITORING PRIORITY – GENERAL SUPERVISION

INDICATOR: The State's SPP/APR includes a State Systemic Improvement Plan (SSIP) that meets the requirements set forth for this indicator.

MEASUREMENT: The State's SPP/APR includes an SSIP that is a comprehensive, ambitious, yet achievable multi-year plan for improving results for children with disabilities. The SSIP includes the components described below.

INSTRUCTIONS FOR THE INDICATOR/MEASUREMENT –

Baseline Data: In its FFY 2013 SPP/APR, due February 1, 2015, the State must provide FFY 2013 baseline data that must be expressed as a percentage and which is aligned with the State-identified Measurable Result(s) for Children with Disabilities.

Targets: In its FFY 2013 SPP/APR, due February 1, 2015, the State must provide measurable and rigorous targets (expressed as percentages) for each of the five years from FFY 2014 through FFY 2018. The State's FFY 2018 target must demonstrate improvement over the State's FFY 2013 baseline data.

Updated Data: In its FFYs 2014 through FFY 2018 SPPs/APRs, due February 2016 through February 2020, the State must provide updated data for that specific FFY (expressed as percentages) and that data must be aligned with the State-identified Measurable Result(s) for Children with Disabilities. In its FFYs 2014 through FFY 2018 SPPs/APRs, the State must report on whether it met its target.

OVERVIEW OF THE THREE PHASES OF THE SSIP: It is of the utmost importance to improve results for children with disabilities by improving educational services, including special education and related services. Stakeholders, including parents of children with disabilities, local educational agencies, the State Advisory Panel, and others, are critical participants in improving results for children with disabilities and should be included in developing, implementing, evaluating, and revising the SSIP and included in establishing the State's targets under Indicator 17. The SSIP should include information about stakeholder involvement in all three phases.

Phase I: Analysis (which the State must include with the February 2, 2015 submission of its SPP/APR for FFY 2013):

- Data Analysis;

- Analysis of State Infrastructure to Support Improvement and Build Capacity;

- State-identified Measurable Result(s) for Children with Disabilities;

- Selection of Coherent Improvement Strategies; and

- Theory of Action.

Phase II: Plan (which, in addition to the Phase 1 content (including any updates) outlined above, the State must include with the February 1, 2016 submission of its SPP/APR for FFY 2014):

- Infrastructure Development;

- Support for local educational agency (LEA) Implementation of Evidence-Based Practices; and

- Evaluation.

Phase III: Implementation and Evaluation (which, in addition to the Phase I and Phase II content (including any updates) outlined above, the State must include with the February 1, 2017 submission of its SPP/APR for FFY 2015, and update in 2018, 2019, and 2020):

- Results of Ongoing Evaluation and Revisions to the SSIP.

SPECIFIC CONTENT OF EACH PHASE OF THE SSIP

Phase I: Analysis

Phase I of the SSIP includes a detailed analysis that will guide the selection of coherent improvement strategies to increase the State's capacity to lead meaningful change in LEAs to improve results for children with disabilities. Phase I must include the following five areas:

- Data Analysis: A description of how the State identified and analyzed key data, including data from SPP/APR indicators, 618 data collections, and other available data as applicable, to: (1) select the State-identified Measurable Result(s) for Children with Disabilities, and (2) identify root causes contributing to low performance.   The description must include information about how the data were disaggregated by multiple variables (e.g., LEA, region, race/ethnicity, gender, disability category, placement, etc.). As part of its data analysis, the State should also consider compliance data and whether those data present potential barriers to improvement. In addition, if the State identifies any concerns about the quality of the data, the description must include how the State will address these concerns. Finally, if additional data are needed, the description should include the methods and timelines to collect and analyze the additional data.

- Analysis of State Infrastructure to Support Improvement and Build Capacity: A description of how the State analyzed the capacity of its current infrastructure to support improvement and build capacity in LEAs to implement, scale up, and sustain the use of evidence-based practices to improve results for children with disabilities. State systems that make up its infrastructure include, at a minimum: governance, fiscal, quality standards, professional development, data, technical assistance, and accountability/monitoring. The description must include current strengths of the systems, the extent the systems are coordinated, and areas for improvement of functioning within and across the systems. The State must also identify current State-level improvement plans and initiatives, including special and general education improvement plans and initiatives, and describe the extent that these initiatives are aligned, and how they are, or could be, integrated with, the SSIP. Finally, the State should identify representatives (e.g., offices, agencies, positions, individuals, and other stakeholders) that were involved in developing Phase I of the SSIP and that will be involved in developing and implementing Phase II of the SSIP.

- State-identified Measurable Result(s) for Children with Disabilities: A statement of the result(s) the State intends to achieve through the implementation of the SSIP. The State-identified result(s) must be aligned to an SPP/APR indicator or a component of an SPP/APR indicator. The State-identified result(s) must be clearly based on the Data and State Infrastructure Analyses and must be a child-level outcome in contrast to a process outcome. The State may select a single result (e.g., increasing the graduation rate for children with disabilities) or a cluster of related results (e.g., increasing the graduation rate and decreasing the dropout rate for children with disabilities).

- Selection of Coherent Improvement Strategies: An explanation of how the improvement strategies were selected, and why they are sound, logical and aligned, and will lead to a measurable improvement in the State-identified result(s). The improvement strategies should include the strategies, identified through the Data and State Infrastructure Analyses, that are needed to improve the State infrastructure and to support LEA implementation of evidence-based practices to improve the State-identified Measurable Result(s) for Children with Disabilities. The State must describe how implementation of the improvement strategies will address identified root causes for low performance and ultimately build LEA capacity to achieve the State-identified Measurable Result(s) for Children with Disabilities.

- Theory of Action: A graphic illustration that shows the rationale of how implementing the coherent set of improvement strategies selected will increase the State's capacity to lead meaningful change in LEAs, and achieve improvement in the State-identified Measurable Result(s) for Children with Disabilities.

Phase II: Plan

The focus of Phase II is on building State capacity to support LEAs with the implementation of evidence-based practices that will lead to measurable improvement in the State-identified Measurable Result(s) for Children with Disabilities. Phase II builds on the data and infrastructure analyses, coherent improvement strategies, and the theory of action developed in Phase I. The plan developed in Phase II includes the activities, steps and resources required to implement the coherent improvement strategies, with attention to the research on

implementation, timelines for implementation and measures needed to evaluate implementation and impact on the State-identified Measurable Result(s) for Children with Disabilities.

- Infrastructure Development: Specify improvements that will be made to the State infrastructure to better support LEAs to implement and scale up evidence-based practices to improve the State-identified Measurable Result(s) for Children with Disabilities. Identify the steps the State will take to further align and leverage current improvement plans and initiatives in the State, including general and special education improvement plans and initiatives, which impact children with disabilities. This section must also identify who will be in charge of implementing the changes to infrastructure, resources needed, expected outcomes, and timelines for completing improvement efforts. In addition, the State should specify how it will involve multiple offices within the State educational agency (SEA), as well as other State agencies, in the improvement of its infrastructure.

- Support for LEA Implementation of Evidence-Based Practices: Specify how the State will support LEAs in implementing the evidence-based practices that will result in changes in LEA, school, and provider practices to achieve the State-identified Measurable Result(s) for Children with Disabilities. This section must identify steps and specific activities needed to implement the coherent improvement strategies, including communication strategies and stakeholder involvement; how identified barriers will be addressed; who will be in charge of implementing; how the activities will be implemented with fidelity; the resources that will be used to implement them; how the expected outcomes of the improvement strategies will be measured; and timelines for completion. In addition, the State should specify how it will involve multiple offices within the SEA (or other State agencies) to support LEAs in scaling up and sustaining the implementation of the evidence-based practices once they have been implemented with fidelity.

- Evaluation: The evaluation must include short-term and long-term objectives to measure implementation of the SSIP and its impact on achieving measurable improvement in State-identified result(s) for children with disabilities. The evaluation must be aligned to the theory of action and other components of the SSIP, include how stakeholders will be involved, and include the methods that the State will use to collect and analyze data to evaluate implementation and outcomes of the SSIP. The evaluation must specify how the State will use the information from the evaluation to examine the effectiveness of the implementation of the SSIP and the progress toward achieving intended improvements in the State-identified Measurable Result(s) for Children with Disabilities, and to make modifications to the SSIP as necessary, and how information from the evaluation will be disseminated to stakeholders.

Phase III: Implementation and Evaluation

In Phase III, the State must, consistent with the evaluation described in Phase II, assess and report on its progress in implementing the SSIP. This will include data and analysis on the extent to which the State has made progress toward and/or met the State-established short-term and long-term objectives for implementation of the SSIP and its progress in achieving the State-identified Measurable Result for Children with Disabilities. If the State intends to continue implementing the SSIP without modifications, the State must describe how the data from the evaluation support this decision. Also, the State must provide a rationale for any revisions that have been made, or revisions the State plans to make, in the SSIP in response to evaluation data, and describe how stakeholders were included in the decision-making process.

Paperwork Burden Statement

Appendix D

ENHANCE Provider Survey

ENHANCE
Promoting quality child outcomes data

Provider Survey

Spring 2012

For more information about ENHANCE, see http://enhance.sri.com

SURVEY ON CHILD OUTCOMES SUMMARY PROCESS
Consent and Introduction

The purpose of this survey is to learn how the process of using the Child Outcomes Summary Form (COSF) is being implemented. This survey is part of a national study designed to improve the child outcomes summary process and the quality of COSF information.

The survey **takes about 15 minutes to complete**. Once you start, you will need to complete the entire survey in one session (you will not be able to save your work and return later to finish it).

Your answers are confidential. No information that identifies you or your individual answers will be shared publicly or with directors or other personnel in your program/district. Findings will be reported using overall responses from the whole group of survey participants. If at least 10 surveys are received from a program/district, those group-level responses to key questions will be shared with administrators for their program/district improvement.

There are no expected risks to participating and your participation is voluntary. You have the right to stop answering questions at any time and there will be no consequences.

As a thank you for taking time to complete the survey, participants will be **entered in a drawing** for a chance to win one of two $100 gift cards. To be included in the drawing, participants must provide a valid email address when prompted.

This survey is being conducted by SRI International, a non-profit research institute working with numerous programs/districts around the country to improve the child outcomes summary process. If you have questions or concerns, you may contact Lauren Barton at 877-697-5765 or email her at enhance@sri.com.

- ❑   I have read the consent information above and agree to participate in this survey now.

This survey uses the term **COSF or child outcomes summary form** to describe the form used to record a rating about the child's functioning on three child outcomes:

- Having positive social relationships,
- Acquiring and using knowledge and skills, and
- Taking appropriate action to meet needs.

The process used to complete the COSF asks people familiar with the child's functioning to combine information from direct assessments, clinical opinion, and family observations to decide on a rating of the child's functioning. Ratings may be labeled from 1-7 or with words such as "completely" to "not at all."

Your program/district may use a different term for the COSF including, but not limited to:

- The ECO form,
- The child outcome questions,
- The child indicator summary form,
- The yellow sheets, or,
- Questions 8, 9, and 10.

Questions in this survey are in multiple choice format. Space is available for additional explanations or comments about the COSF at the end of the survey.

SURVEY ON CHILD OUTCOMES SUMMARY PROCESS
Section 1: About Your Training and Experience with the COSF

1. Currently, at your program/district, approximately how many COSFs have you participated in? *(Count all COSFs where you had any involvement in identifying the rating. Examples include: discussing the rating with others, facilitating discussions with others, and/or completing the form yourself.)*

- ❑ Zero
- ❑ 1-10
- ❑ 11-30
- ❑ 31-50
- ❑ More than 50

2. Have you received information or training about the Child Outcomes Summary Form (COSF) process?

- ❑ Yes
- ❑ No

What training or information have you received? *(Check all that apply.)*

- ❑ In-person state level training event
- ❑ In-person local or regional training event
- ❑ Online or video training module
- ❑ Webinar or training conference call
- ❑ Review of COSF training materials
- ❑ One-on-one training
- ❑ Ongoing feedback from a supervisor or program director
- ❑ Website resources (e.g., ECO Center or state website)
- ❑ I provide training on the COSF to others
- ❑ Other

Please describe "Other":

_____
_____

3. How many total hours have you spent being trained or learning about the COSF process? *(Give your best estimate.)*

- ❑ None
- ❑ Less than 1 hour
- ❑ 1-2 hours
- ❑ 3-4 hours
- ❑ 5-8 hours
- ❑ 9-15 hours
- ❑ More than 15 hours

4. On average, how long does it take to identify a child's outcome ratings and provide documentation on the form?

In your estimate, please include:
- time to identify the rating.
- time discussing the child's functioning if it is directly related to the rating decision or exceeds discussions about child's functioning that would have occurred anyway.
- time to complete information on the form.

Do <u>not</u> include:
- time for data entry of the form, if this is done after completion.

- ❑ 1-15 minutes
- ❑ 16-30 minutes
- ❑ 31-45 minutes
- ❑ 46-60 minutes
- ❑ More than 60 minutes

5. Have you ever used the decision tree?
- ❑ Yes
- ❑ No
- ❑ Don't know what it is

6. How helpful is the decision tree in reaching a rating?
- ❑ Very helpful
- ❑ Helpful
- ❑ Not helpful
- ❑ Not at all helpful
- ❑ I can't judge, I have seen it, but never used it
- ❑ I can't judge, I have not seen it before or I don't know what it is

7. Please rate HOW TRUE the following statements are: *(Check one in each row.)*

	Very True	Mostly True	Somewhat True	A Little True	Not at all True
a. I understand the meaning of each of the three outcomes.	❑	❑	❑	❑	❑
b. I understand how to apply the criteria for each of the 7 rating points.	❑	❑	❑	❑	❑
c. I understand the difference between functional behaviors and discrete skills.	❑	❑	❑	❑	❑
d. I understand what is age-expected functioning in each of the three outcome areas.	❑	❑	❑	❑	❑
e. I understand why we are collecting child outcomes data.	❑	❑	❑	❑	❑
f. I understand what happens with the child outcomes data that we collect.	❑	❑	❑	❑	❑

SURVEY ON CHILD OUTCOMES SUMMARY PROCESS
Section 2: About Knowledge and Skills Related to the Child Outcomes Summary Process

8. Please rate HOW TRUE the following statements are: *(Check one in each row.)*

	Very True	Mostly True	Somewhat True	A Little True	Not at all True
a. I know how to explain the need for the child outcomes ratings to families.	❑	❑	❑	❑	❑
b. I know how to discuss the child's functioning in the three outcome areas with others who know the child.	❑	❑	❑	❑	❑
c. I know how to identify how the child uses his/her skills to perform meaningful, everyday tasks.	❑	❑	❑	❑	❑
d. I know how to collect information about the child's functioning across settings and situations.	❑	❑	❑	❑	❑
e. I know how to compare the child's functioning to age-expected functioning.	❑	❑	❑	❑	❑
f. I know how to talk with families about age-expected functioning.	❑	❑	❑	❑	❑
g. I know how to identify whether or not the child made any progress in the outcome areas (needed at exit or for follow up discussions).	❑	❑	❑	❑	❑

Section 3: About Your Experience with the COSF

9. In HOW MANY of your COSFs have you experienced the following in your current program/district? *(Check one in each row.)*

	All of the children's COSFs (100%)	Most of the children's COSFs (76-99%)	Many of the children's COSFs (51-75%)	Some of the children's COSFs (26-50%)	A few of the children's COSFs (1-25%)	None of the children's COSFs (0%)
a. The family provided input about the child's functioning.	❑	❑	❑	❑	❑	❑
b. The rating was decided by a team that included at least one other professional and me.	❑	❑	❑	❑	❑	❑
c. Information about the child's functioning from multiple settings and situations was used in deciding the ratings.	❑	❑	❑	❑	❑	❑
d. I was **not** involved in deciding the ratings, but I provided input on the child's functioning.	❑	❑	❑	❑	❑	❑
e. The family was present during the decision of the child outcomes ratings.	❑	❑	❑	❑	❑	❑
f. At least one other professional in addition to me provided input about the child's functioning.	❑	❑	❑	❑	❑	❑
g. Information from one or more assessment tools was used in deciding the ratings.	❑	❑	❑	❑	❑	❑
h. All involved considered information carefully in order to identify an accurate rating.	❑	❑	❑	❑	❑	❑
i. There was enough information about the child's functioning in each outcome to decide on a rating.	❑	❑	❑	❑	❑	❑
j. There was enough time to review the child's functioning in each of the three outcome areas.	❑	❑	❑	❑	❑	❑
k. I was confident that the ratings given were accurate.	❑	❑	❑	❑	❑	❑
l. The process used for deciding ratings matched my understanding of how it is supposed to be done.	❑	❑	❑	❑	❑	❑
m. The ratings were selected to make the program look good.	❑	❑	❑	❑	❑	❑

For the following questions, consider all professionals involved in the rating decisions since you have been at your *current* program/district. Include both program/district staff and any contracted providers who participate in the COSF process.

9b. How many other professionals involved in COSF ratings understand…
(Check one in each row.)

	All	Almost All	Many	Some	A Few	None
a. The meaning of each of the three outcomes.	❑	❑	❑	❑	❑	❑
b. How to apply the criteria for each of the 7 rating points.	❑	❑	❑	❑	❑	❑
c. The difference between functional behaviors and discrete skills.	❑	❑	❑	❑	❑	❑
d. What is age-expected functioning in each of the three outcome areas.	❑	❑	❑	❑	❑	❑

SURVEY ON CHILD OUTCOMES SUMMARY PROCESS

Section 4: About Experiences with the Child Outcomes Summary Process

10. Please rate HOW TRUE the following statements are: *(Check one in each row.)*

	Very True	Mostly True	Somewhat True	A Little True	Not at all True
a. Information from assessment tools we use is very helpful in determining ratings for the three outcomes.	❑	❑	❑	❑	❑
b. I receive feedback from someone such as a supervisor on the child outcomes summary ratings or the form.	❑	❑	❑	❑	❑
c. Ratings tend to be low at entry relative to the child's actual level of functioning.	❑	❑	❑	❑	❑
d. There's too much additional paperwork associated with the child outcomes summary process.	❑	❑	❑	❑	❑
e. The child outcomes summary process is a good way to collect data on child outcomes.	❑	❑	❑	❑	❑
f. The ratings given are higher than the child's actual level of functioning.	❑	❑	❑	❑	❑
g. Child outcomes summary ratings are too subjective.	❑	❑	❑	❑	❑

	Very True	Mostly True	Somewhat True	A Little True	Not at all True
h. Ongoing support related to the child outcomes summary process is adequate.	❑	❑	❑	❑	❑
i. The child outcomes summary process is a useless activity.	❑	❑	❑	❑	❑
j. The ratings given are lower than the child's actual level of functioning.	❑	❑	❑	❑	❑
k. Ratings are more accurate when parents are present for the rating decision.	❑	❑	❑	❑	❑
l. It is difficult for individuals involved in identifying child outcomes ratings to reach consensus on one or more of the three outcomes.	❑	❑	❑	❑	❑
m. Ratings tend to be high at exit relative to the child's actual level of functioning.	❑	❑	❑	❑	❑
n. I receive helpful feedback about the child outcomes summary form.	❑	❑	❑	❑	❑
o. Ratings are less accurate when parents are present for the rating decision.	❑	❑	❑	❑	❑
p. I like the three outcomes.	❑	❑	❑	❑	❑
q. The child outcomes summary process emphasizes age-expected functioning too much.	❑	❑	❑	❑	❑

SURVEY ON CHILD OUTCOMES SUMMARY PROCESS

Section 5: About program/district activities

11. Someone in our current program/district…. *(Check one in each row.)*

	Yes	No	Don't Know
a. checks child outcome summary forms after they are completed to ensure the ratings are accurate.	❑	❑	❑
b. provides feedback to those who are involved in the COSF process.	❑	❑	❑
c. is available to provide me with ongoing support if I ask for it.	❑	❑	❑
d. trains professionals new to the child outcomes summary process.	❑	❑	❑

Section 6: Impact of the Child Outcomes Summary Process on Practice/Services

12. The child outcomes summary process… *(Check one in each row.)*

	Very True	Mostly True	Somewhat True	A Little True	Not at all True
a. makes me more aware of children's functioning relative to expectations for their age.	❑	❑	❑	❑	❑
b. negatively impacts the assessment process.	❑	❑	❑	❑	❑
c. leads to better IFSP or IEP outcomes.	❑	❑	❑	❑	❑
d. improves the way we work as a team.	❑	❑	❑	❑	❑
e. has negative impacts on my relationships with families.	❑	❑	❑	❑	❑
f. helps me think about children's functioning across settings and with different people.	❑	❑	❑	❑	❑
g. helps me focus on functional use of skills to perform meaningful tasks instead of discrete skills.	❑	❑	❑	❑	❑
h. leads to poorer quality IFSP or IEP outcomes.	❑	❑	❑	❑	❑
i. helps me think about children's progress over time.	❑	❑	❑	❑	❑
j. improves the quality of my conversations with families about their child.	❑	❑	❑	❑	❑
k. takes time away from other important activities.	❑	❑	❑	❑	❑
l. improves the assessment process.	❑	❑	❑	❑	❑
m. helps focus discussion on the "whole child."	❑	❑	❑	❑	❑

13. Have there been any other POSITIVE impacts on your practice that are not included in the questions above?

❑ Yes
❑ No

Please describe:
_____
_____

14. Have there been any other NEGATIVE impacts on your practice that are not included in the questions above?

❑ Yes
❑ No

Please describe:
_____
_____

15. Overall, what has been the impact of the child outcomes summary process on your work with children and families?

❑ Very Positive
❑ Positive
❑ Neutral
❑ Negative
❑ Very Negative

Section 7: About You

16. Which of the following describe your professional role in your program?
(Check all that apply.)
- ❑ Early Interventionist / Child Development Specialist / Infant Specialist / Developmental Therapist
- ❑ Special Education teacher
- ❑ General Education Teacher
- ❑ Service Coordinator / Case Manager
- ❑ Family Resource Coordinator
- ❑ Speech-Language Pathologist
- ❑ Speech-Language Pathologist Assistant
- ❑ Occupational Therapist
- ❑ Occupational Therapy Assistant
- ❑ Physical Therapist
- ❑ Physical Therapy Assistant
- ❑ Psychologist
- ❑ Educational Diagnostician / Evaluator / Examiner
- ❑ Social worker / Counselor
- ❑ Nutritionist
- ❑ Director / Administrator
- ❑ COSF trainer
- ❑ Data Entry Clerk
- ❑ Other

Please describe "Other":
_____
_____
_____
_____

17. Of the children you work with in a typical month, what percent are in the following age groups? *(Please count work with families based on the age of child in the family. Enter "0" if none.)*

Birth to 3 years (%)_____
3 through 5 years (%)_____
Other ages (%)_____
Total (must equal 100%)_____

18. How long have you been providing services to young children with disabilities? *(Working with children with disabilities under 6 years of age.)*
- ❑ Less than 1 year
- ❑ 1-2 years
- ❑ 3-5 years
- ❑ 6-10 years
- ❑ 11 years or more

19. Have you worked (in any capacity) with young children birth to five *without* disabilities? *(e.g., child care, teaching, assessment)*
- ❑ Yes
- ❑ No

For how long?
- ❑ Less than 1 year
- ❑ 1-2 years
- ❑ 3-5 years
- ❑ 6 years or more

20. How old are you?
- ❑ Under 30
- ❑ 30-39
- ❑ 40-49
- ❑ 50-59
- ❑ 60-69
- ❑ 70 or above

21. What is your gender?
- ❑ Male
- ❑ Female

22. Which of these describes you? *(Check all that apply.)*
- ❑ African-American/Black
- ❑ Asian-American
- ❑ Hispanic/Latino
- ❑ Caucasian/White
- ❑ American Indian
- ❑ Pacific Islander
- ❑ Do not wish to disclose
- ❑ Other

Please describe "Other":

_____
_____

In which state are you located?

_____

What is your program name?

_____

Do you work with: *(Check all that apply.)*
- ❑ Early Intervention (EI)
- ❑ Early Childhood Special Education (ECSE)

SURVEY ON CHILD OUTCOMES SUMMARY PROCESS
Section 8: About Your Ideas to Improve the COSF Process and Other Comments

23. Is there anything else you want to tell us about the child outcomes summary process or this survey?

_____
_____

Please provide the following information about yourself.

Name:_____

Email address:_____

(This information will be confidential. All survey respondents providing email addresses will be entered in a drawing for a chance to win one of two $100 gift cards. Your email address also will help us avoid any duplications in survey data and avoid us sending you unnecessary reminders to complete the survey. If you do not wish to provide an email address, simply type in abc@abc.com.)

Thank you for your time!
If you have any questions, please contact or call 877-697-5765.
We will enter your name/email address in the incentive drawing and let you know if you win a $100 gift card.

This concludes the survey.

Appendix E

Provider Survey Data for EI and ECSE

Note: Appendix exhibit numbers relate to exhibit numbers in the text.

Exhibit E-3.3   Characteristics and Training of EI and ECSE Providers in Study 1

Characteristics of EI Providers	*n*	%	Characteristics of ECSE Providers	*n*	%	
Age group served (*n* = 554)			**Age group served (*n* = 384)**			
EI	472	85	ECSE	302	79	
Both EI and ECSE	82	15	Both EI and ECSE	82	21	
Role EI and Both (*n* = 534)			**Role ECSE and Both (*n* = 378)**			
Early interventionist/teacher	249	47	Early interventionist/teacher	187	50	
Therapists	192	36	Therapists	163	43	
Other	93	17	Other	28	7	
Years of experience EI and Both (*n* = 529)			**Years of experience ECSE and Both (*n* = 377)**			
2 years or less	50	10	2 years or less	31	8	
3 to 5 years	87	16	3 to 5 years	57	15	
6 years or more	392	74	6 years or more	289	77	
Number of COS ratings made EI and Both (*n* = 554)			**Number of COS ratings made ECSE and Both (*n* = 384)**			
10 or less	109	20	10 or less	93	24	p < .001
11 to 30	107	19	11 to 30	161	42	
31 or more	338	61	31 or more	130	34	
Hours of COS training EI and Both (*n* = 554)			**Hours of COS training ECSE and Both (*n* = 384)**			
None	28	5	None	19	5	p < .001
Some, but less than 4 hours	401	72	Some, but less than 4 hours	253	66	
5 to 8 hours	77	14	5 to 8 hours	72	19	
9 hours or more	48	9	9 hours or more	40	10	
Amount of time to complete COS ratings (*n* = 554)			**Amount of time to complete COS ratings (*n* = 384)**			
1 – 15 minutes	256	46	1 – 15 minutes	82	21	
16 – 30 minutes	189	34	16 – 30 minutes	140	36	
> 30 minutes	109	20	> 30 minutes	162	42	

E-2

Exhibit E-3.4   EI and ECSE Providers' Self-Reported Training and Feedback on COS Process Provided by Their Programs

EI Providers (N=537)

	Yes	No	Don't know
Program trains professionals new to the child outcomes summary process	66%	5%	30%
Program checks child outcome summary forms after they are completed to ensure ratings are accurate	34%	10%	56%
Program is available to provide me with ongoing support if I ask for it	81%	3%	16%
Program provides feedback to those who are who are involved in the COS process	47%	21%	32%

■ Yes ■ No ■ Don't know

ECSE Providers (N=381)

	Yes	No	Don't know	
Program trains professionals new to the child outcomes summary process	71%	6%	23%	p < .01
Program checks child outcome summary forms after they are completed to ensure ratings are accurate	41%	9%	50%	p < .04
Program is available to provide me with ongoing support if I ask for it	83%	5%	12%	
Program provides feedback to those who are who are involved in the COS process	52%	22%	27%	

■ Yes ■ No ■ Don't know

E-3

Exhibit E-3.5   EI and ECSE Providers' Perceptions of Quality of Training and Feedback about COS Process

EI Providers

I receive feedback from someone such as a supervisor on the child outcomes summary ratings or the form (N = 541)
- Very or mostly true: 29%
- Somewhat or a little true: 37%
- Not at all true: 34%

I receive helpful feedback about the child outcomes summary form (N = 537)
- Very or mostly true: 21%
- Somewhat or a little true: 45%
- Not at all true: 34%

Ongoing support related to the child outcomes summary process is adequate (N = 537)
- Very or mostly true: 47%
- Somewhat or a little true: 44%
- Not at all true: 9%

■ Very or mostly true ■ Somewhat or a little true ■ Not at all true

ECSE Providers

I receive feedback from someone such as a supervisor on the child outcomes summary ratings or the form (N = 382)
- Very or mostly true: 40%
- Somewhat or a little true: 31%
- Not at all true: 29%

p < .001

I receive helpful feedback about the child outcomes summary form (N = 381)
- Very or mostly true: 22%
- Somewhat or a little true: 42%
- Not at all true: 36%

Ongoing support related to the child outcomes summary process is adequate (N = 381)
- Very or mostly true: 48%
- Somewhat or a little true: 43%
- Not at all true: 9%

■ Very or mostly true ■ Somewhat or a little true ■ Not at all true

Exhibit E-3.6   EI and ECSE Providers' Self-Reported Understanding of COS Process

EI Providers (N=553)

	Very or mostly true	Somewhat or a little true	Not at all true
I understand the meaning of each of the 3 outcomes	85%	14%	1%
I understand how to apply the criteria for each of the 7 rating points	78%	20%	2%
I understand the difference between functional behaviors and discrete skills	74%	23%	3%
I understand what is age-expected functioning in each of the three outcome areas	89%	10%	1%
I understand why we are collecting child outcomes data	66%	28%	6%
I understand what happens with the child outcomes data that we collect	37%	46%	17%

ECSE Providers (N=384)

	Very or mostly true	Somewhat or a little true	Not at all true	
I understand the meaning of each of the 3 outcomes	86%	13%	1%	
I understand how to apply the criteria for each of the 7 rating points	82%	17%	1%	p < .01
I understand the difference between functional behaviors and discrete skills	70%	26%	3%	
I understand what is age-expected functioning in each of the three outcome areas	90%	10%		
I understand why we are collecting child outcomes data	66%	28%	6%	
I understand what happens with the child outcomes data that we collect	40%	43%	17%	

■ Very or mostly true ■ Somewhat or a little true ■ Not at all true

E-5

Exhibit E-3.7   EI and ECSE Providers' Self-Reported Skills in Making COS Rating

EI Providers (N=552)

Statement	Very or mostly true	Somewhat or a little true	Not at all true
I know how to explain the need for the child outcomes ratings to families	54%	37%	9%
I know how to discuss the child's functioning in the three outcomes areas with others who know the child	83%	14%	2%
I know how to identify how the child uses his/her skills to perform meaningful, everyday tasks	91%	9%	
I know how to collect information about the child's functioning across settings and situations	89%	10%	1%
I know how to compare the child's functioning to age-expected functioning	92%	7%	1%
I know how to talk with families about age-expected functioning	92%	7%	1%
I know how to identify whether or not the child made any progress in the outcome areas (needed at exit or for follow up discussions)	89%	10%	2%

ECSE Providers (N=384)

Statement	Very or mostly true	Somewhat or a little true	Not at all true
I know how to explain the need for the child outcomes ratings to families	60%	32%	8%
I know how to discuss the child's functioning in the three outcomes areas with others who know the child	83%	16%	2%
I know how to identify how the child uses his/her skills to perform meaningful, everyday tasks	90%	9%	1%
I know how to collect information about the child's functioning across settings and situations	91%	8%	1%
I know how to compare the child's functioning to age-expected functioning	92%	8%	
I know how to talk with families about age-expected functioning	90%	9%	1%
I know how to identify whether or not the child made any progress in the outcome areas (needed at exit or for follow up discussions)	89%	10%	1%

■ Very or mostly true   ■ Somewhat or a little true   ■ Not at all true

All not significant.

E-6

Exhibit E-3.8   EI and ECSE Providers' Perceptions of Other Team Members' Understanding of COS Process

EI Providers (N=521)

The meaning of each of the 3 outcomes	71% · 18% · 8% · 1%	2%
How to apply the criteria for each of the 7 rating points	61% · 26% · 10% · 0%	3%
The difference between functional behaviors and discrete skills	62% · 22% · 12% · 1%	3%
What is age-expected functioning in each of the three outcome areas	77% · 16% · 6% · 1%	1%

0% 20% 40% 60% 80% 100%

■All or almost all   □Many   ■Some   □A few   ■None

ECSE Providers (N=366)

The meaning of each of the 3 outcomes	66% · 19% · 12% · 0%	3%
How to apply the criteria for each of the 7 rating points	62% · 23% · 13%	3%
The difference between functional behaviors and discrete skills	54% · 26% · 16%	5%
What is age-expected functioning in each of the three outcome areas	68% · 22% · 9%	2%

0% 20% 40% 60% 80% 100%

p < .001

■All or almost all   □Many   ■Some   □A few   ■None

E-7

Exhibit E-3.9 EI and ECSE Providers' Report about Their Experiences Using the COS Process: Who Participates in COS Ratings

EI Providers (N=548)

The family was present during the decision of the child outcomes ratings
42% | 6% | 5% | 15% | 32%

The family provided input about the child's functioning
65% | 14% | 6% | 7% | 9%

The rating was decided by a team that included at least one other professional and me
77% | 9% | 5% | 5% | 4%

At least one other professional in addition to me provided input about the child's functioning
82% | 7% | 4% | 3% | 4%

I was not involved in deciding the ratings, but I provided input on the child's functioning
11% | 7% | 7% | 24% | 52%

0% 20% 40% 60% 80% 100%

■ All or almost all (76-100%)   □ Many (51-75%)
■ Some (26-50%)   □ A few (1-25%)
■ None (0%)

ECSE Providers (N=383)

The family was present during the decision of the child outcomes ratings
28% | 5% | 4% | 8% | 55%   p < .001

The family provided input about the child's functioning
72% | 11% | 7% | 6% | 6%

The rating was decided by a team that included at least one other professional and me
73% | 9% | 7% | 7% | 4%

At least one other professional in addition to me provided input about the child's functioning
73% | 8% | 7% | 7% | 4%   p < .004

I was not involved in deciding the ratings, but I provided input on the child's functioning
9% | 4% | 7% | 16% | 65%   p < .001

0% 20% 40% 60% 80% 100%

■ All or almost all (76-100%)   □ Many (51-75%)
■ Some (26-50%)   □ A few (1-25%)
■ None (0%)

Exhibit E-3.10 EI and ECSE Providers' Report about Their Experiences Using COS Process: Information Used to Make COS Rating

EI Providers

Information about the child's functioning from multiple settings and situations was used in deciding the ratings (N = 548)	74% 14% 6% 4% 3%
Information from one or more assessment tools was used in deciding the ratings (N = 548)	74% 10% 5% 6% 5%
All involved considered information carefully in order to identify an accurate rating (N = 542)	83% 10% 3% 2% 2%
There was enough information about the child's functioning in each outcome to decide on a rating (N = 542)	86% 9% 3% 1% 2%

ECSE Providers

Information about the child's functioning from multiple settings and situations was used in deciding the ratings (N = 383)	78% 8% 6% 5% 3%
Information from one or more assessment tools was used in deciding the ratings (N = 383)	77% 8% 6% 4% 5%
All involved considered information carefully in order to identify an accurate rating (N = 382)	87% 5% 3% 2% 3% p < .02
There was enough information about the child's functioning in each outcome to decide on a rating (N = 382)	84% 9% 5% 1% 2%

Legend (both charts):
- All or almost all (76-100%)
- Many (51-75%)
- Some (26-50%)
- A few (1-25%)
- None (0%)

Exhibit E-3.11 EI and ECSE Providers' Report about Their Experiences Using COS Process: Perceptions of Integrity of COS Rating Process

EI Providers (N=542)

There was enough time to review the child's functioning in each of the three outcome areas	80% / 11% / 5% / 2% / 2%
I was confident that the ratings given were accurate	74% / 16% / 5% / 2% / 3%
The process used for deciding ratings matched my understanding of how it is supposed to be done	76% / 15% / 4% / 2% / 4%
The ratings were selected to make the program look good	5% / 2% / 4% / 3% / 87%

Scale: 0% — 20% — 40% — 60% — 80% — 100%

Legend:
- ■ All or almost all (76-100%)
- □ Many (51-75%)
- ■ Some (26-50%)
- □ A few (1-25%)
- ■ None (0%)

ECSE Providers (N=382)

There was enough time to review the child's functioning in each of the three outcome areas	80% / 9% / 6% / 2% / 4%
I was confident that the ratings given were accurate	78% / 13% / 5% / 1% / 3%
The process used for deciding ratings matched my understanding of how it is supposed to be done	81% / 11% / 3% / 2% / 4%
The ratings were selected to make the program look good	6% / 2% / 2% / 1% / 89%

Scale: 0% — 20% — 40% — 60% — 80% — 100%

Legend:
- ■ All or almost all (76-100%)
- □ Many (51-75%)
- ■ Some (26-50%)
- □ A few (1-25%)
- ■ None (0%)

All not significant.

Exhibit E-3.12 EI and ECSE Providers' Self-Reported Attitudes about COS Process

EI Providers

Statement	Very or mostly true	Somewhat or a little true	Not at all true
Information from assessment tools we use is very helpful in determining ratings for the three outcomes (N = 541)	77%	21%	2%
I like the three outcomes (N = 537)	26%	51%	23%
The child outcomes summary process is a good way to collect data on child outcomes (N = 541)	30%	52%	18%
There's too much additional paperwork associated with the child outcomes summary process (N = 541)	27%	49%	24%
The child outcomes summary process is a useless activity (N = 537)	20%	50%	30%
The child outcomes summary process emphasizes age-expected functioning too much (N = 537)	11%	47%	42%

ECSE Providers

Statement	Very or mostly true	Somewhat or a little true	Not at all true
Information from assessment tools we use is very helpful in determining ratings for the three outcomes (N = 382)	78%	20%	2%
I like the three outcomes (N = 381)	26%	53%	22%
The child outcomes summary process is a good way to collect data on child outcomes (N = 381)	31%	55%	14%
There's too much additional paperwork associated with the child outcomes summary process (N = 381)	39%	49%	12%
The child outcomes summary process is a useless activity (N = 381)	23%	50%	28%
The child outcomes summary process emphasizes age-expected functioning too much (N = 381)	8%	48%	44%

p < .001

■ Very or mostly true ■ Somewhat or a little true ■ Not at all true

Exhibit E-3.13 EI and ECSE Providers' Perceptions about Quality of COS Rating Process

EI Providers

Statement	Very or mostly true	Somewhat or a little true	Not at all true
Ratings tend to be low at entry relative to the child's actual level of functioning (N = 541)	16%	36%	48%
The ratings given are higher than the child's actual level of functioning (N = 541)	%	30%	67%
The ratings given are lower than the child's actual level of functioning (N = 537)	%	28%	69%
Ratings tend to be high at exit relative to the child's actual level of functioning (N = 537)	8%	30%	62%
Ratings are more accurate when parents are present for the rating decision (N = 537)	26%	39%	36%
Ratings are less accurate when parents are present for the rating decision (N = 537)	11%	33%	56%
Child outcomes summary ratings are too subjective (N = 541)	36%	53%	11%
It is difficult for individuals involved in identifying child outcomes ratings to reach consensus on one or more of the three outcomes (N = 537)	5%	48%	47%

■ Very or mostly true ■ Somewhat or a little true ■ Not at all true

ECSE Providers (N=381)

Statement	Very or mostly true	Somewhat or a little true	Not at all true	
Ratings tend to be low at entry relative to the child's actual level of functioning	22%	37%	42%	
The ratings given are higher than the child's actual level of functioning	%	28%	68%	
The ratings given are lower than the child's actual level of functioning	%	31%	67%	
Ratings tend to be high at exit relative to the child's actual level of functioning	9%	28%	63%	
Ratings are more accurate when parents are present for the rating decision	17%	43%	40%	p < .003
Ratings are less accurate when parents are present for the rating decision	8%	36%	57%	
Child outcomes summary ratings are too subjective	29%	56%	14%	p < .03
It is difficult for individuals involved in identifying child outcomes ratings to reach consensus on one or more of the three outcomes	5%	47%	48%	

■ Very or mostly true ■ Somewhat or a little true ■ Not at all true

Exhibit E-3.14 EI and ECSE Providers' Perceptions about Positive Impacts of COS Process on their Practice

EI Providers (N=535)

Statement	Very or mostly true	Somewhat or a little true	Not at all true
Makes me more aware of children's functioning relative to expectations for their age	35%	42%	22%
Leads to better IFSP or IEP outcomes	18%	36%	46%
Improves the way we work as a team	20%	35%	45%
Helps me think about children's functioning across settings and with different people	37%	43%	20%
Helps me focus on functional use of skills to perform meaningful tasks instead of discrete skills	34%	40%	27%
Helps me think about children's progress over time	42%	39%	19%
Improves the quality of my conversations with families about their child	22%	32%	46%
Improves the assessment process	18%	37%	45%
Helps focus discussion on the "whole child"	31%	40%	29%

ECSE Providers (N=378)

Statement	Very or mostly true	Somewhat or a little true	Not at all true	
Makes me more aware of children's functioning relative to expectations for their age	37%	38%	26%	
Leads to better IFSP or IEP outcomes	15%	41%	44%	p < .01
Improves the way we work as a team	20%	39%	41%	
Helps me think about children's functioning across settings and with different people	38%	43%	19%	
Helps me focus on functional use of skills to perform meaningful tasks instead of discrete skills	32%	40%	28%	
Helps me think about children's progress over time	38%	42%	20%	
Improves the quality of my conversations with families about their child	20%	34%	46%	
Improves the assessment process	16%	44%	40%	
Helps focus discussion on the "whole child"	31%	44%	25%	

Legend: ■ Very or mostly true ■ Somewhat or a little true ■ Not at all true

E-13

Exhibit E-3.15 EI and ECSE Providers' Perceptions about Negative Impacts of COS Process on their Practice

EI Providers (N=535)

Negatively impacts the assessment process	4% 25%	71%
Has negative impacts on my relationships with families	2% 15%	83%
Leads to poorer quality IFSP or IEP outcomes	3% 9%	88%
Takes time away from other important activities	25% 42%	32%

0%   20%   40%   60%   80%   100%

■ Very or mostly true   ■ Somewhat or a little true   ■ Not at all true

ECSE Providers (N=378)

Negatively impacts the assessment process	2% 20%	78%	p < .02
Has negative impacts on my relationships with families	7%	92%	p < .0003
Leads to poorer quality IFSP or IEP outcomes	2% 7%	91%	
Takes time away from other important activities	39% 40%	21%	p < .0001

0%   20%   40%   60%   80%   100%

■ Very or mostly true   ■ Somewhat or a little true   ■ Not at all true

E-14

Appendix F

Psychometric Characteristics of the BDI-2 and Vineland-II

Exhibit F-1   Psychometric Characteristics of the Battelle Developmental Inventory, second edition (BDI-2) and the Vineland Adaptive Behavior Scales, second edition (Vineland-II)

Exhibit F-1   Psychometric Characteristics of the Battelle Developmental Inventory, second edition (BDI-2) and the Vineland Adaptive Behavior Scales, second edition (Vineland-II)

	Vineland-II			BDI-2		
Reliability	**Subdomains**	**Domains**	**ABS - Adaptive Behavior Composite**	**Subdomains**	**Domains**	**Total Score**
Internal consistency	.69-.96 [90% of subdomains > .75]	.79-.95 [91.7% of domains >.85]	.95-.98 Mean = .97	.85-.95	.90-.96	.98-.99 Mean = .99
	Split-half method, across children 0-5			Split-half method, across children 0-7, detail by age shows also true for children 0-5 years		
Test-retest reliability	Mean *r* 0-2 yrs = .86 3-6 yrs = .84	Mean *r* 0-2 yrs = .91 3-6 yrs = .89	Mean *r* 0-2 yrs = .96 3-6 yrs = .94	Correlations: 2 yrs, *r* = .79-.90 4 yrs, *r* = .74-.91	Correlations: 2 yrs, *r* = .87-.90 4 yrs, *r* = .87-.92	Correlations: 2 yrs, *r* = .93 4 yrs, *r* = .94
	14-30 days apart			2-25 days apart – 80% within 2 weeks		
Rater/scorer agreement	Mean *r* 0-6 yrs = .77	Mean *r* 0-6 yrs = .74	Mean *r* 0-6 yrs = .80	Inter-scorer agreement. For fine and perceptual motor items: 97-99% agreement between scores on 3 sets of items, original and new examiners		
	Interrater reliability. 0-24 days, 2 caregivers - usually both parents					
Validity						
Developmental acquisition of skills and behaviors	Means and standard deviations across ages (including 0-5) are provided. For subdomains, domains, and the composite score, means and standard deviations are consistent with established developmental patterns and rates of growth.			In tryout and standardization samples, conducted item discrimination (including item-total *r*), percentage correct at successive age levels, model-data-fit statistics from the Rasch model, and differential item functioning analyses with gender and racial/ethnic groups. Items with poor fit statistics caused by poor discrimination, poor age sensitivity, or lack of contribution to domain structure were deleted from final version of BDI-2, yielding a strong final version.		
Factor structure	For children 3-6 yrs: A four factor solution fit the data well (CFI = .96), with strong path coefficients as expected. The four factor solution fit the data significantly better than a 1 factor solution (chi square difference = 254.5 (df = 4) *p* < .001).					
Expected differences in clinical samples	Scores on subdomains, domains, and the composite score showed expected differences between a samples of verbal and nonverbal children 3-6 with autism and their non-clinically diagnosed, same-aged peers.			Scores on domains, and the total score showed expected differences between a samples children for each of the following compared to their non-diagnosed same-aged peers: with autism (2-7 yrs), with cognitive delays (2-7 yrs), with developmental delays (2 mo-6 yrs), with motor delays (2 mo-6.5 yrs), those who had premature births (1 mo-3 yrs), and speech and language delays (1-7 yrs).		
Relationships to other measures	Vineland-II domains with: Vineland ABS 0-2 yrs, *r* =.65-.91, 3-6 yrs, *r* = .85-.94; composite ABAS-II[1] 0-5 yrs, *r* = .49-.65; relevant subdomains BASC-II[2] 3-5 yrs, *r* = .36-.62. Vineland-II ABS with: Vineland ABS 0-2 yrs, *r* = .82, 3-6 yrs, *r* = .91; ABAS-II composite 0-5 yrs, *r* = .70; BASC-II 3-5 yrs, *r* = .46.			BDI-2 domains with: BDI total score, *r* = .51-.66, BSID-II[3], *r* = .61-.75 and PLS-4, *r* = .63-.73; low correlations found with less relevant domains. BDI-2 total score with: BDI total score, *r* = .78.		

[1] Comparing Vineland-II subdomain scores with the general adaptive composite score on the Adaptive Behavior Assessment System, 2nd Edition.
[2] Comparing relevant subdomains to each other on the Behavior Assessment System for Children, 2nd Edition.
[3] Comparing relevant subdomains to each other on Bayley Scales of Infant Development, 2nd Edition (BSID-II) and Preschool Language Scales, 4th Edition (PLS-4).

Appendix G

Crosswalks of BDI-2 and Vineland-II to Child Outcomes

Introduction to the ECO "Crosswalks" of Birth-to-Five Assessment Instruments to Early Childhood Outcomes

Battelle Developmental Inventory—Second Edition (BDI-2)
Table 1. Crosswalk of Domain Items to Child Outcomes

Summary Information: Vineland Adaptive Behavior Scales Second Edition (Vineland-II)

Introduction to the
ECO "Crosswalks" of Birth-to-Five Assessment Instruments
to Early Childhood Outcomes

The "crosswalks" identify relationships between assessment instruments and the three child outcomes on which state Part C and 619 programs must report to the Office of Special Education Programs (OSEP). States must report children's progress in the outcome areas of:

1. Positive social emotional skills (including positive social relationships)
2. Acquisition and use of knowledge and skills (including early language/communication and early literacy)
3. Use of appropriate behaviors to meet their needs

Understanding the Three Child Outcomes
The three child outcomes reflect a global, overarching goal for all children: to be active and successful participants now and in the future, in a variety of settings. Accomplishments in various dimensions of each outcome area move a child toward that goal, as follows.

Outcome 1: Positive social emotional skills involves relating with adults, relating with other children, and, for older children, following rules related to groups or interacting with others. This outcome includes attachment/separation/autonomy, expressing emotions and feelings, learning rules and expectations, and social interactions and play.

Outcome 2: Acquisition and use of knowledge and skills involves thinking, reasoning, remembering, problem solving, using symbols and language, and understanding physical and social worlds. This outcome includes early concepts (symbols, pictures, numbers, classification, spatial relationships), imitation, object permanence, and language skills.

Outcome 3. Use of appropriate behaviors to meet their needs involves taking care of basic needs, getting from place to place, using tools and, for older children, contributing to their own health and safety. This outcome includes integrating motor skills to complete tasks, self-help skills (dressing, feeding, grooming, toileting, household responsibility), and acting on the world to get what one wants.

The three child outcomes are **functional** in that they reflect a child's ability to take meaningful action in the context of everyday living. The outcome areas cross developmental domains, emphasizing the integration of skills and behaviors across domains for meaningful action. The presence of an isolated skill or behavior gives limited information about a child's functioning. The outcomes address whether a child can integrate skills and put them to use across settings and situations. A child's natural use of pointing to indicate what he needs or wants, for example, reflects functioning better than his ability to point to objects when asked to do so by a tester as part of an assessment.

Assessing the Three Child Outcomes

Assessing children's functioning in the three outcome areas requires multiple sources of information, including observation, family input, and data from one or more assessment tool. Observation and family input provide information about children's functioning across situations and settings. Data from the administration of a commercial assessment tool can be used to compare a child's skills and behaviors to those of his same-age peers. A limitation in the use of currently available assessment tools, however, is that they are not designed for direct measurement of the three outcomes. Most are organized around domains, with items separated into discrete areas of development, such as expressive language, receptive language, cognitive, gross motor, and fine motor. Some call for standardized assessment items to be administered in a setting other than the child's natural environment, making it difficult to use the information to determine whether a child uses this skill in everyday life.

The review of currently available assessment tools has been a primary activity for states as they consider options for measuring child outcomes. A key question is 'how much information will an assessment tool provide about the attainment of the three outcomes?' ECO developed the crosswalks to assist states and programs in making decisions about what instruments might be useful to include in an outcomes measurement system, including reporting to OSEP. The crosswalks indicate how the contents of the assessments maps to the three outcomes. The crosswalks also allow comparisons across instruments to see their various strengths and weaknesses with regard to the three outcomes.

Crosswalking Purposes

The ECO crosswalks display how content on a given assessment instrument is related to each of the three child outcomes. Organized in a table format with assessment areas assigned to each outcome, they provide a visual depiction of coverage. By showing how an assessment tool covers each of the three outcome areas, the crosswalks are meant to help states, programs, and providers see the extent of information available in an outcome area from a given assessment tool. Some of the crosswalks include examples of assessment items to illustrate the types of skills and behaviors the tool targets. In addition to comparing tools, states can use the crosswalks to determine areas in which additional information will need to be collected, such as through observation and family input, to make up for any shortcomings in the data provided by an assessment tool.

The crosswalks are not meant to be used as a "checklist" or "score sheet" for measuring child outcomes. ECO does not recommend the use of isolated items or areas of items from any given tool. We support the use of assessment instruments in the way in which they were designed to be used. In addition, given the functional nature of the outcomes, we support the use of assessment tools in conjunction with other sources of information about a child's functioning, such as observation and family report.

Crosswalks were generated for instruments based on the frequency of informal requests from states. Priority was also given to instruments that states identified for outcomes measurement in the State Performance Plans submitted to the Office of Special Education Programs in 2005. These crosswalks are presented as a service to the field. **The ECO Center does not endorse the use of any specific assessment instrument.** Thus, a completed crosswalk does not constitute the endorsement of an instrument. If a crosswalk of an instrument is not available it is because, given the reality of finite

resources, it has not yet been completed. For more information about crosswalk content or specific instruments, please email a request to staff@the-eco-center.org.

Crosswalking Processses
We use the following guidelines in completing the crosswalks.

Level at which assessment tools are crosswalked. Criterion-referenced or curriculum-based assessment tools are typically crosswalked at the sub-area level, using the developer's headings. As appropriate, examples of items from a sub-area are included to illustrate the aspects of development that relate to the outcome. Norm-referenced tests[1] always are crosswalked at the lowest level that the tool developers recommend valid interpretation of the data and have provided normative information. This is usually at a subscale or sub-domain level,

Assignment of assessment area or sub-area to an outcome. We place areas/sub-areas/items from each assessment tool under the outcome to which they are most closely linked conceptually. For example, items about getting along with peers go with Outcome 1. Decisions are based on content of the area rather than the heading title because headings do not always reflect the range of behaviors and skills included. Particularly in the sub-areas of language and learning, it is difficult to assign items to outcomes when item content lacks specificity. In such cases we assume that the item pertains to a general, overarching acquisition and use of knowledge and skills, and therefore make the assignment to Outcome 2.

Double classification. Although many sub-areas or items can be double classified because of the interrelated nature of development in young children, we try to minimize double classification in order to minimize redundancy. Sub-areas or items that relate to a second outcome area, but not as strongly as they relate to a primary outcome area, are only classified with the primary area. Sub-areas or items are double classified when it is felt that the information contributes equally or nearly equally toward understanding achievement of more than one outcome.

"Precursor skills." Some items on assessment tools target skills that, while not functional in and of themselves, may lead to functional behaviors. For example, a child's ability to use a pincer grasp may lead to his ability to feed himself or hold a pencil. Prescursor skills that are clearly linked to one of the outcomes are placed with that outcome. We assign general or cross-cutting precursor skills to Outcome 2, as part of general acquisition and use of knowledge and skills. We also note in the crosswalks when precursor skills for functional behaviors skills, such as those associated with motor development, may not be appropriate or expected for children with sensory, motor, or other impairments.

Inclusion of every skill in a crosswalk. Not all skills in an assessment tool can be classified. Items/areas are left out that do not contribute to understanding the child's functional abilities in any particular outcome area. The decision not to classify areas such as sleeping, riding a tricycle, or moving to music is not meant to imply that such experiences are not important for young children.

---

[1] Crosswalks of norm-referenced instruments include a note providing information about the lowest appropriate threshold for crosswalking on that specific instrument.

Status of the Crosswalks
Crosswalks are available on the ECO website in draft form. On each crosswalk, a footnote indicates the date and the status of the draft. For instance, some drafts are preliminary; others have been revised based on input from assessment tool authors or publishers. Please compare the date on any crosswalk you are using to the version on the web site to see if you have the latest version since revisions are frequently posted.

Questions and comments are encouraged and should be sent to staff@the-eco-center.org. Also, please contact us if you are interested in a crosswalk that is not found on our website. Additional information about measuring outcomes can be found on our website at www.the-eco-center.org.

Battelle Developmental Inventory—Second Edition (BDI-2)
Table 1. Crosswalk of Domain Items to Child Outcomes

When using the BDI-2 as one of multiple sources of information for measuring child outcomes, the following crosswalk applies. As a standardized, norm-referenced assessment tool, the BDI-2's domain scores are the smallest unit of information that can be used if scores are converted directly to measure child outcomes. However, as one of multiple sources of information, items from BDI-2 domains can contribute to the understanding of a child's functioning in each of the three outcome areas. This table shows how various items from domains map to the three child outcomes. If converting domain scores directly to child outcomes, see Table 2 of this document.

	Outcome 1 Positive social relationships	Outcome 2 Acquires and use skills and knowledge	Outcome 3 Takes action to meet needs
Domain: **ADAPTIVE**			**Self-care** ▪ SC1-35 (feeding, dressing, toileting) **Personal responsibility** ▪ PR1-25 (movement, safety, organizes own activities)
Domain: **PERSONAL-SOCIAL**	**Adult interaction** ▪ AI1-30 (looks at, responds to adults, initiates social contact) **Peer interaction** ▪ PI1-25 (responds to, plays with other children, shares properly, plays cooperatively) **Self-concept and social role** ▪ SR1-45 (prescursors to self awareness and self awareness, describes own feelings)		

Note: Draft developed by the Early Childhood Outcomes (ECO) Center and revised based on preliminary feedback from users and the tool publisher and/or developers. The draft may be subject to further changes. We welcome your feedback to <staff@the-eco-center.org>.   1

Summary Information: Vineland Adaptive Behavior Scales Second Edition (Vineland-II)

Name	Vineland Adaptive Behavior Scales Second Edition (Vineland-II)
Publisher	Pearson Assessments
Website for information	http://ags.pearsonassessments.com
Cost	Vineland-II Survey Forms & Teacher Rating Form Starter Set: $210.99 Includes Vineland-II Survey Forms Manual, Teacher Rating Form Manual and 10 of each of the following forms: Survey Interview Form, Parent/Caregiver Rating Form, Survey Forms Report to Parents, Survey Forms Report to Caregivers, Teacher Rating Form and Teacher Rating Form Report to Parents and Caregivers
Age range:	Survey Interview Form, Parent/Caregiver Rating Form, Expanded Interview Form—0 through 90; Teacher Rating Form—3 through 21-11
Purpose	The Vineland II measures personal and social skills from birth to adulthood, with multiple purposes: • Support diagnosis of mental retardation, autism, and developmental delays • Determine eligibility or qualification for special services • Plan rehabilitation or intervention programs • Track and report progress
Areas included	Five domains: Communication, daily living skills, socialization, motor, and maladaptive (optional)
Time to administer	Survey Interview and Parent/Caregiver Rating Forms- 20-60 minutes
Scored	
Type of Scores	Domain and Adaptive Behavior Composite—Standard scores (M = 100, SD = 15), percentile ranks, adaptive levels, age equivalents; Subdomain—V-scale score (M = 15, SD = 3), Adaptive levels, age equivalents; Survey Interview, Parent/Caregiver Rating Form, Expanded Interview Form—V-scale score, maladaptive levels for the optional Maladaptive Behavior Index
Age norms	Yes.

Age ranges given for items	
How frequently it can be given	not specified
Standardized tasks	
Based on observation in natural settings	
Based on information requested from parents or providers	Yes. Survey Interview Form is semi-structured parent interview and Parent/ Caregiver Rating Form completed by parents. Teacher Rating Form also available.
Data provided on reliability	
Data provided on validity	
Web-based data entry	no.
Electronic scoring	Software available for use with the Survey Interview Form, Parent/ Caregiver Rating Form, and Teacher Rating Form. Software scores and has options for multiple report formats.
Other languages	Survey Interview Form Record Booklets, Survey Form Report to Parents, & Survey Form Report to caregivers available in Spanish
Who administers	Level C; Vineland-II test users should have a PH.D. in psychology or be a certified or licensed school psychologist or social worker.
Training available through the publisher	

Vineland Adaptive Behavior Scales Second Edition (Vineland-II): Crosswalk to Child Outcomes
Note: Because the Vineland is a norm-referenced, standardized assessment, the subdomain scores are the smallest unit of information that can be used to reach conclusions about the extent to which a child is demonstrating each of the functional outcomes. This table shows how the Vineland's subdomains map to the three outcomes. Under each subdomain, the X indicates the outcome area to which the subdomain score contributes information. The item information under the X provides the rationale for why the subdomain was classified as providing information for that outcome.

Domain/Subdomain	Outcome 1 Positive social relationships	Outcome 2 Acquire and use skills and knowledge	Outcome 3 Takes action to meet needs
Communication/ Expressive	X	X	X
	Interactive speech	Pre-speech expression, beginning to talk, interactive speech, expressing complex ideas	Expresses wants or needs
Communication/ Receptive		X	
		Understanding, listening and attending, following instructions	
Communication/ Written		X	
		Beginning to read, reading skills, writing skills	
Daily Living Skills/ Personal			X
			Eating and drinking, toileting, grooming, dressing, bathing, health care

Daily Living Skills/ Community			X
			Telephone skills, restaurant skills, job skills, computer skills
Daily Living Skills/ Domestic			X
			Safety at home, kitchen chores, housekeeping
Socialization/ Interpersonal Relationships	X		
	Expressing emotions		
Socialization/ Play and Leisure Time	X		
	Playing, sharing and cooperating, going places with friends		
Socialization/ Coping Skills	X		X
	Controlling impulses, apologizing, transitions, responsibility, manners		Appropriate social caution
Motor Skills			X
			Drawing and using scissors, manipulating objects, using keyboard

Note: Draft developed at state request. This draft has not been through the ECO review process to establish consistency with the crosswalks posted on the ECO web site.

	Outcome 1 **Positive social relationships**	Outcome 2 **Acquires and use skills and knowledge**	Outcome 3 **Takes action to meet needs**
Domain: COMMUNICATION	**Receptive** ▪ RC5-8; 27 (responds to person, converses) **Expressive** ▪ EC13, 18 (communicates with others) ▪ EC30-31 (follows conventional rules of conversation)	**Receptive** ▪ RC1-4; 5-9 (precursors to understanding language ▪ RC9-26 (associates words with objects, actions; recalls events from a story) ▪ RC28-40 (understands grammar, identifies sounds in words) **Expressive** ▪ EC1-8; 10-11, 13 (precursors to using language) ▪ EC12, 14-15, 17, 19-21 (uses words) ▪ EC22-29 (asks questions, uses sentences) ▪ EC32-37 (uses grammar) ▪ EC40-45 (uses words, gives directions)	**Expressive** ▪ EC9; 16 (gestures, uses words to indicate wants, needs) ▪ EC38 (communicates feelings)
Domain: MOTOR			**Gross** ▪ GM1-45 (movement precursors to taking action)* **Fine** ▪ FM1-30 (movement precursors to taking action)* **Perceptual** ▪ PM 1-25 (movement precursors to taking action)*

Note: Draft developed by the Early Childhood Outcomes (ECO) Center and revised based on preliminary feedback from users and the tool publisher and/or developers. The draft may be subject to further changes. We welcome your feedback to <staff@the-eco-center.org>.

2

	Outcome 1 Positive social relationships	Outcome 2 Acquires and use skills and knowledge	Outcome 3 Takes action to meet needs
Domain: COGNITIVE		**Attention and memory** ▪ AM1-30 (precursors to acquisition and use of skills and knowledge) **Reasoning and academic skills** ▪ RA1-35 (recognizes cause, matches colors, gives objects on requests, completes opposite analogies, sequences events, rote counts, write letters, solves math problems) **Perception and concepts** ▪ PC1-40 (explores, imitates, matches, sorts, identifies shapes and objects, understands time, categorizes objects by function, matches simple words, groups objects, sorts by multiple properties)	

*This domain includes precursor skills for functional behaviors, which may not be appropriate or expected for some children, including those with sensory, motor, or other impairments.

3

Battelle Developmental Inventory—Second Edition (BDI-2)
Table 2. Crosswalk of Domain Scores to Child Outcomes

If converting scores directly to child outcomes, the following crosswalk applies. As a standardized, norm-referenced assessment tool, the BDI-2's domain scores are the smallest unit of information that can be used. This table shows how the 5 domains map to the three child outcomes. For each domain, an X indicates the outcome area to which the *majority* of items contributes information.

	Outcome 1 Positive social relationships	Outcome 2 Acquires and uses skills and knowledge	Outcome 3 Takes action to meet needs
Domain: ADAPTIVE ▪ Self-care ▪ Personal responsibility			X
Domain: PERSONAL-SOCIAL ▪ Adult interaction ▪ Peer interaction ▪ Self-concept and social role	X		
Domain: COMMUNICATION ▪ Receptive ▪ Expressive		X	

Note: Draft developed by the Early Childhood Outcomes (ECO) Center and revised based on preliminary feedback from users and the tool publisher and/or developers. The draft may be subject to further changes. We welcome your feedback to <staff@the-eco-center.org>.

4

	Outcome 1 Positive social relationships	Outcome 2 Acquires and uses skills and knowledge	Outcome 3 Takes action to meet needs
Domain: **MOTOR** ▪ Gross ▪ Fine ▪ Perceptual			**X***
Domain: **COGNITIVE** ▪ Attention and memory ▪ Reasoning and academic skills ▪ Perception and concepts		**X**	

*This domain includes precursor skills for functional behaviors, which may not be appropriate or expected for some children, including those with sensory, motor, or other impairments.

Note: Draft developed by the Early Childhood Outcomes (ECO) Center and revised based on preliminary feedback from users and the tool publisher and/or developers. The draft may be subject to further changes. We welcome your feedback to <staff@the-eco-center.org>. 5

Appendix H

ENHANCE Child and Family Information Form

ENHANCE
Promoting quality child outcomes data

Child and Family Information Form

Date completed *(mm/dd/yy)*: _____ / _____ / _____

Provider *(person completing the form) (print)*:   First: _____   Last: _____

Email: _____   *Phone (with area code):*   _____ - _____ - _____

Child's Name *(print)*:   First: _____   Last: _____

Primary Parent/Caregiver *(print)*:

First: _____   *Last:* _____

Address/Apt: _____

City/State/Zip: _____

Home Phone: _____ - _____ - _____

Cell Phone: _____ - _____ - _____

Work Phone: _____ - _____ - _____

Best time to call: _____

Email: _____

Alternate Parent/Caregiver *(print)*:

First: _____   *Last:* _____

Address/Apt: _____

City/State/Zip: _____

Home Phone: _____ - _____ - _____

Cell Phone: _____ - _____ - _____

Work Phone: _____ - _____ - _____

Best time to call: _____

Email: _____

Child's Date of Birth *(mm/dd/yy)*: _____ / _____ / _____   **Child's Gender:** ____ *Male*   ____ *Female*

Languages other than English spoken in the home:
    ____ *None*
    ____ *Spanish*
    ____ *Other (specify:_____)*

Child's race/ethnicity *(check all that apply)*:
    ____ *African American*
    ____ *Asian-American*
    ____ *Caucasian/White*
    ____ *Hispanic/Latino*
    ____ *Other (specify:_____)*

Highest level of education completed
by Primary Parent/Caregiver:

    ____ *Less than high school*
    ____ *High school diploma or GED*
    ____ *Vocational/technical training or some college*
    ____ *College degree*
    ____ *Graduate training*

Child's disability type: _____

Child's disability category:
For ages 0 to 3:   ____ *Developmental Delay*
                     ____ *Diagnosed/Established Condition*

For ages 3 – 5:   ____ *Developmental Delay*
                     ____ *Speech-Language Delay*
                     ____ *Other (specify:_____)*

Date of first IFSP/IEP in this program *(mm/dd/yy)*: _____ / _____ / _____
Date of first COSF in this program *(mm/dd/yy)*:   _____ / _____ / _____
Anticipated date of exit from Part C or Part B Preschool services *(mm/dd/yy)*:   _____ / _____ / _____

When completed: *FAX directly to 877-364-2620, or*
                Mail to ENHANCE Study, SRI International, 333 Ravenswood Ave. BS188, Menlo Park, CA 94025, or
                G*ive to local ENHANCE Program/District Liaison.*

For questions: *Email ENHANCE@sri.com or call toll free 877-697-5765. For project info, go to http://ENHANCE.sri.com.*

Appendix I

Additional Descriptive Information about Longitudinal Sample

Exhibit I-1. Average COS Ratings Based on Type of Disability in Longitudinal Sample at Program Entry and Exit

Longitudinal Sample (*n* = 70)	Positive Social Relationships	Knowledge and Skills	Taking Action to Meet Needs
Program Entry	Mean (SD)	Mean (SD)	Mean (SD)
Diagnosed Condition (EI, *n* = 12)	4.6 (2.4)	4.2 (2.2)	3.9 (2.2)
Developmental Delay (EI, *n* = 34)	5.0 (1.2)	4.1 (1.4)	4.6 (1.3)
Developmental Delay (ECSE, *n* = 8)	4.4 (1.4)	4.6 (0.7)	4.5 (1.2)
Speech-Language Impaired (ECSE, *n* = 11)	5.4 (1.5)	4.8 (1.6)	5.5 (1.8)
Other condition (ECSE, *n* = 5)	2.2 (.04)	2.2 (0.4)	2.4 (0.5)
Program Exit	Mean (SD)	Mean (SD)	Mean (SD)
Diagnosed Condition (EI, *n* = 12)	5.1 (2.2)	4.8 (2.1)	4.5 (2.0)
Developmental Delay (EI, *n* = 34)	5.9 (1.1)	5.6 (1.2)	5.7 (1.3)
Developmental Delay (ECSE, n =8)	5.4 (1.3)	5.1 (0.8)	5.5 (1.1)
Speech-Language Impaired (ECSE, *n* =11)	5.6 (1.7)	6.1 (1.5)	6.3 (1.6)
Other condition (ECSE, *n* = 5)	4.2 (1.1)	4.6 (0.9)	4.8 (1.6)

Exhibit I-2. Variations in Mean COS Ratings Based on Descriptive Characteristics of the Longitudinal Sample

	Longitudinal Sample (*n* = 70)					
	Positive Social Relationships		Knowledge and Skills		Action to Meet Needs	
	Entry Mean (SD)	Exit Mean (SD)	Entry Mean (SD)	Exit Mean (SD)	Entry Mean (SD)	Exit Mean (SD)
Gender						
Male (*n* = 38)	4.7 (1.6)	5.5 (1.3)	4.3 (1.6)	5.4 (1.3)	4.5 (1.7)	5.6 (1.3)
Female (*n* = 32)	4.7 (1.7)	5.6 (1.8)	4 (1.6)	5.5 (1.6)	4.4 (1.6)	5.4 (1.8)
Age						
<1 year (*n* = 17)	4.6 (1.9)	5.6 (1.5)	4 (1.8)	5.3 (1.5)	3.9 (1.9)	5.1 (1.6)
1 year (*n* = 11)	4.7 (1.5)	5.3 (1.6)	4.2 (1.7)	5.4 (1.7)	3.9 (1.4)	5.3 (1.8)
2 years (*n* = 18)	5.3 (1.3)	5.9 (1.4)	4.2 (1.4)	5.6 (1.4)	5.3 (0.8)	5.8 (1.5)
3 years (*n*= 12)	4 (2)	5.3 (1.7)	4 (1.7)	5.5 (1.6)	4.2 (2)	5.8 (1.6)
>= 4 years (*n* = 12)	4.75 (1.5)	5.2 (1.4)	4.3 (1.43)	5.4 (1.1)	4.9 (1.6)	5.6 (1.4)
Time in between entry and exit COS						
< 9 months (*n* = 16)	5.3 (1.3)	5.4 (1.4)	4.3 (1.2)	5.2 (1.2)	5 (1.4)	5.7 (1.3)
9-11 months (*n* = 19)	4.7 (1.5)	5.8 (1.8)	4.3 (1.6)	5.6 (1.6)	4.7 (1.4)	5.4 (1.8)
1 to 2 years (*n* = 18)	5.2 (1.1)	5.8 (0.9)	4.5 (1.4)	5.8 (1.1)	4.7 (1.5)	5.8 (1.2)
2 years or more (*n* = 17)	3.8 (2.2)	5 (1.8)	3.5 (2)	5.1 (1.7)	3.5 (2)	5.1 (1.9)

Appendix J

Child Assessment Study Analyses: Relationships between Child and Team Characteristics and Entry COS Ratings on Each Outcome

Throughout this appendix we report findings from a series of regression analyses to investigate the extent to which specific characteristics are related to COS ratings on each of the three outcomes at entry. In each regression, either the total score of the ABILITIES Index or the relevant subarea of the ABILITIES Index is included as a covariate. All ABILITIES Index scores have been reversed such that higher values represent higher levels of functioning.

Throughout this appendix we report findings from a series of regression analyses to investigate the extent to which specific child and team characteristics are related to COS ratings on each of the three outcomes in the child assessments study Entry Sample ($n = 153$). These analyses consider the following outcomes: early intervention program, male gender, race/ethnicity, child age, family member presence, number of COS team members, and service coordinator, as well as a model showing these predictors in a combined model. Results from these analyses are summarized in the main report in Section 4 (Study 2).

- Each regression model was run the following two different ways:
  - With the ABILITIES Index total score (reversed) as a covariate, taking into account all aspects of the child's level of functioning
  - With the ABILITIES Index subarea[1] most relevant to the child outcome area as the covariate (i.e., Social/Communication for positive social relationships; Cognitive/Communication for knowledge and skills; and Structural Integrity for taking action to meet needs). Subarea scores also are reversed so that higher scores represent higher levels of functioning.
- Results are organized by the outcome area being predicted. Regressions are shown in pairs, with the ABILITIES Index total score as a covariate being reported first and displayed with an **orange** figure title and the regression with the ABILTIES Index subarea as a covariate reported next and a figure title in **blue**.

---

[1]   More detail about the development of these ABILTIES Index subareas is contained in the main report.

1. Positive Social Relationships

Exhibit 1.1a **Regression of Early Intervention Predicting COS Ratings for Positive Social Relationships in Entry Sample (_n_ = 153) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Positive Social Relationships		
	_B_	_SE B_	_β_
Intercept	-3.52***	1.02	0.00
ABILITIES Index	0.08***	0.01	0.52
Early Intervention	0.49*	0.23	0.15
R^2		0.31	
_F_		33.68***	

*_p_ < .05, **_p_ < .01, ***_p_ < .001.

Exhibit 1.1b **Regression of Early Intervention Predicting COS Ratings for Positive Social Relationships in Entry Sample (_n_ = 153) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Positive Social Relationships		
	_B_	_SE B_	_β_
Intercept	0.52	0.41	0.00
ABILITIES— Social/ Communication	0.23***	0.02	0.68
Early Intervention	-0.32	0.23	-0.10
R^2		0.41	
_F_		52.36***	

*_p_ < .05, **_p_ < .01, ***_p_ < .001.

Exhibit 1.2a **Regression of Male Gender Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 153) (ABILITIES Index as Covariate)**

	COS Entry Ratings: Positive Social Relationships		
Variable	*B*	*SE B*	*β*
Intercept	-3.41**	1.04	0.00
ABILITIES Index	0.08***	0.01	0.54
Male Gender	-0.02	0.28	-0.01
R^2		0.29	
F		30.44***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.2b **Regression of Male Gender Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 153) (ABILITIES Index Subarea as Covariate)**

	COS Entry Ratings: Positive Social Relationships		
Variable	*B*	*SE B*	*β*
Intercept	0.26	0.45	0.00
ABILITIES — Social/ Communication	0.22***	0.022	0.65
Male Gender	0.35	0.21	0.11
R^2		0.41	
F		53.05***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.3a   Regression of Race/Ethnicity Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 153) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	$SE\ B$	β
Intercept	-3.12**	1.03	0.00
ABILITIES Index	0.08***	0.01	0.52
Race/Ethnicity			
Hispanic	0.38	0.32	0.08
Black	-0.61*	0.29	-0.14
Other	-0.09	0.49	-0.01
R^2		0.32	
F		17.36***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.3b   Regression of Race/Ethnicity Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 153) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	$SE\ B$	β
Intercept	0.65	0.44	0.00
ABILITIES— Social/ Communication	0.21***	0.02	0.62
Race/Ethnicity			
Hispanic	0.39	0.30	0.08
Black	-0.19	0.28	-0.04
Other	0.19	0.46	0.03
R^2		0.41	
F		26.16***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.4a   Regression of Child Age at Entry Predicting COS Ratings for Positive Social Relationships in Entry Sample (*n* = 153) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	-3.05**	1.04	0.00
ABILITIES Index	0.08***	0.01	0.53
Child Age at Entry (months)	-0.01	0.01	-0.12
R²		0.30	
F		32.49***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.4b   Regression of Child Age at Entry Predicting COS Ratings for Positive Social Relationships in Entry Sample (*n* = 153) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	-0.54	0.55	0.00
ABILITIES — Social/ Communication	0.25***	0.02	0.73
Child Age at Entry (months)	0.02**	0.01	0.21
R²		0.44	
F		58.15***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.5a Regression of Family Member Presence at the COS Team Meeting Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 153) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	-3.19**	1.039	0.00
ABILITIES Index	0.08***	0.01	0.52
Parent or family member at the COS Team Meeting	0.23	0.23	0.07
R²		0.29	
F		29.0***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.5b Regression of Family Member Presence at the COS Team Meeting Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 153) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	0.57	0.42	0.00
ABILITIES— Social/ Communication	0.21***	0.02	0.62
Parent or family member at the COS Team Meeting	0.14	0.21	0.04
R²		0.40	
F		47.35***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.6a   **Regression of Number of COS Team Members Predicting COS Ratings for Positive Social Relationships in Entry Sample (*n* = 153) (ABILITIES Index as Covariate)**

	COS Entry Ratings: Positive Social Relationships		
Variable	*B*	*SE B*	*β*
Intercept	-3.42**	1.09	0.00
ABILITIES Index	0.08***	0.01	0.54
Number of COS Team Members	0.03	0.06	0.04
R²		0.29	
F		28.52***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.6b   **Regression of Number of COS Team Members Predicting COS Ratings for Positive Social Relationships in Entry Sample (*n* = 153) (ABILITIES Index Subarea as Covariate)**

	COS Entry Ratings: Positive Social Relationships		
Variable	*B*	*SE B*	*β*
Intercept	0.56	0.49	0.00
ABILITIES — Social/ Communication	0.22***	0.02	0.63
Number of COS Team Members	0.01	0.06	0.01
R²		0.39	
F		47.04***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.7a **Regression of Service Coordinator on COS Team Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 153) (ABILITIES Index as Covariate)**

	COS Entry Ratings: Positive Social Relationships		
Variable	B	$SE\ B$	β
Intercept	-4.17***	1.09	0.00
ABILITIES Index	0.08***	0.01	0.55
Service Coordinator on COS Team	0.62*	0.31	0.14
R^2		0.31	
F		33.26***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.7b **Regression of Service Coordinator on COS Team Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 153) (ABILITIES Index Subarea as Covariate)**

	COS Entry Ratings: Positive Social Relationships		
Variable	B	$SE\ B$	β
Intercept	0.55	0.46	0.00
ABILITIES— Social/ Communication	0.22***	0.02	0.63
Service Coordinator on COS Team	0.05	0.28	0.01
R^2		0.40	
F		50.75***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.8a   Regression of X on COS Team Predicting COS Ratings for Positive Social Relationships in Entry Sample (_n_ = 153) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	_B_	_SE B_	_β_
Intercept	-3.72**	1.23	0.00
ABILITIES Index	0.07***	0.01	0.48
Early Intervention	0.95*	0.44	0.29
Male Gender	0.09	0.24	0.03
Ethnicity			
Hispanic	0.56	0.33	0.12
Black	-0.71*	0.30	-0.17
Other	0.17	0.50	0.02
Child Age at Entry (months)	0.01	0.01	0.12
Parent at the COS Team Meeting	0.35	0.27	0.11
Number of COS Team Members	0.06	0.07	0.07
R^2		0.37	
F		8.64***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Exhibit 1.8b   Regression of X on COS Team Predicting COS Ratings for Positive Social Relationships in Entry Sample (*n* = 153) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	-1.84*	0.83	0.00
ABILITIES— Social/ Communication	0.24***	0.03	0.70
Early Intervention	0.91*	0.39	0.27
Male Gender	0.37	0.22	0.11
Race/Ethnicity			
Hispanic	0.57	0.30	0.12
Black	-0.12	0.29	-0.03
Other	0.17	0.46	0.02
Child Age at Entry (months)	0.04***	0.01	0.41
Parent at the COS Team Meeting	0.18	0.25	0.06
Number of COS Team Members	0.02	0.07	0.02
R^2		0.48	
F		13.59***	

*p < .05, **p < .01, ***p < .001.

Note: Service coordinator participation was correlated significantly with EI program status and so was removed from the multiple regression model.

2. Knowledge and Skills

Exhibit 2.1a   **Regression of Early Intervention Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 153) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	-3.29**	0.98	0.00
ABILITIES Index	0.07***	0.01	0.53
Early Intervention	0.00	0.22	0.00
*R*²		0.28	
F		28.96***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.1a   **Regression of Early Intervention Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 153) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	0.74*	0.36	0.00
ABILITIES— Cognitive/Communication	0.28***	0.03	0.67
Early Intervention	-0.56**	0.21	-0.18
*R*²		0.41	
F		51.12***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.2a   Regression of Male Gender Predicting COS Ratings for Knowledge and Skills in Entry Sample ($n = 153$) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	**SE B**	**β**
Intercept	-3.32**	0.99	0.00
ABILITIES Index	0.07***	0.01	0.53
Male Gender	0.06	0.22	0.02
R^2		0.28	
F		29.01***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Exhibit 2.2b   Regression of Male Gender Predicting COS Ratings for Knowledge and Skills in Entry Sample ($n = 153$) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	**SE B**	**β**
Intercept	0.47	0.40	0.00
ABILITIES—Cognitive/Communication	0.26***	0.03	0.63
Male Gender	0.31	0.20	0.10
R^2		0.39	
F		47.41***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Exhibit 2.3a   Regression of Race/Ethnicity Predicting COS Ratings for Knowledge and Skills in Entry Sample ($n = 153$) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	SE B	β
Intercept	-2.93**	0.97	0.00
ABILITIES Index	0.07***	0.01	0.51
Race/Ethnicity			
Hispanic	0.05	0.30	0.01
Black	-0.86**	0.28	-0.21
Other	-0.47	0.47	-0.07
R^2		0.33	
F		17.94***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Exhibit 2.3b   Regression of Race/Ethnicity Predicting COS Ratings for Knowledge and Skills in Entry Sample ($n = 153$) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	SE B	β
Intercept	0.942*	0.39	0.00
ABILITIES— Cognitive/Communication	0.24***	0.02	0.59
Race/Ethnicity			
Hispanic	0.17	0.29	0.04
Black	-0.47	0.27	-0.12
Other	-0.19	0.44	-0.03
R^2		0.39	
F		24.01***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Exhibit 2.4a   Regression of Child Age at Entry Predicting COS Ratings for Knowledge and Skills in Entry Sample (n = 153) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	$SE\ B$	β
Intercept	-3.10**	0.99	0.00
ABILITIES Index	0.07***	0.01	0.53
Child Age at Entry (months)	-0.01	0.01	-0.06
R^2		0.28	
F		29.53***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.4b   Regression of Child Age at Entry Predicting COS Ratings for Knowledge and Skills in Entry Sample (n = 153) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	$SE\ B$	β
Intercept	-0.18	0.48	0.00
ABILITIES— Cognitive/Communication	0.29***	0.03	0.69
Child Age at Entry (months)	0.02**	0.01	0.20
R^2		0.41	
F		52.09***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.4a Regression of Family member presence at the COS Team Meeting Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 153) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	-3.21**	0.98	0.00
ABILITIES Index	0.07***	0.01	0.52
Parent or family member at the COS Team Meeting	0.38	0.22	0.13
R^2		0.30	
F		30.66***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.5b Regression of Family member presence at the COS Team Meeting Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 153) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	0.45	0.38	0.00
ABILITIES—Cognitive/Communication	0.26***	0.03	0.62
Parent or family member at the COS Team Meeting	0.32	0.19	0.10
R^2		0.42	
F		50.94***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.6a   Regression of Number of COS Team Members Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 153) (ABILITIES Index as Covariate)

	COS Entry Ratings: Knowledge and Skills		
Variable	*B*	*SE B*	*β*
Intercept	-3.39**	1.04	0.00
ABILITIES Index	0.07***	0.01	0.54
Number of COS Team Members	0.02	0.06	0.03
R²		0.29	
F		28.55***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.6b   Regression of Number of COS Team Members Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 153) (ABILITIES Index Subarea as Covariate)

	COS Entry Ratings: Knowledge and Skills		
Variable	*B*	*SE B*	*β*
Intercept	0.64	0.43	0.00
ABILITIES—Cognitive/Communication	0.27***	0.03	0.64
Number of COS Team Members	-0.03	0.05	-0.04
R²		0.41	
F		49.03***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.7a   Regression of Service Coordinator on COS Team Predicting COS Ratings for Knowledge and Skills in Entry Sample (n = 153) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	$SE\ B$	β
Intercept	-3.88***	1.04	0.00
ABILITIES Index	0.08***	0.01	0.54
Service Coordinator on COS Team	0.48	0.29	0.11
R^2		0.29	
F		30.82***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.7b   Regression of Service Coordinator on COS Team Predicting COS Ratings for Knowledge and Skills in Entry Sample (n = 153) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	$SE\ B$	β
Intercept	0.77	0.41	0.00
ABILITIES—Cognitive/Communication	0.25***	0.03	0.62
Service Coordinator on COS Team	-0.07	0.28	-0.02
R^2		0.38	
F		45.48***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.8a   Regression of X on COS Team Predicting COS Ratings for Knowledge and Skills in Entry Sample (n = 153) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	-2.17	1.15	0.00
ABILITIES Index	0.07***	0.01	0.48
Early Intervention	-0.26	0.41	-0.08
Male Gender	0.20	0.22	0.06
Race/Ethnicity			
Hispanic	0.12	0.31	0.03
Black	-0.97***	0.28	-0.24
Other	-0.45	0.47	-0.07
Child Age at Entry (months)	-0.01	0.01	-0.14
Parent at the COS Team Meeting	0.68**	0.25	0.22
Number of COS Team Members	-0.05	0.07	-0.06
R^2		0.38	
F		9.19***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Exhibit 2.8b   Regression of X on COS Team Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 153) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	-0.14	0.75	0.00
ABILITIES—Cognitive/Communication	0.28	0.03	0.66
Early Intervention	0.05***	0.37	0.02
Male Gender	0.39	0.20	0.12
Race/Ethnicity			
Hispanic	0.27	0.28	0.06
Black	-0.42	0.27	-0.11
Other	-0.36	0.42	-0.05
Child Age at Entry (months)	0.02	0.01	0.19
Parent at the COS Team Meeting	0.63**	0.23	0.21
Number of COS Team Members	-0.12	0.06	-0.14
R^2		0.49	
F		14.57***	

*p < .05, **p < .01, ***p < .001.

Note: Service coordinator participation was correlated significantly with EI program status and so was removed from the multiple regression model.

3. Action to Meet Needs

Exhibit 3.1a **Regression of Early Intervention Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 153) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-3.77***	0.92	0.00
ABILITIES Index	0.08***	0.01	0.60
Early Intervention	-0.34	0.20	-0.11
R^2		0.36	
F		41.75***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.1b **Regression of Early Intervention Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 153) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-0.79	0.75	0.00
ABILITIES— Structural Integrity	0.10***	0.01	0.51
Early Intervention	0.04	0.22	0.01
R^2		0.26	
F		26.4***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.2a **Regression of Male Gender Predicting COS Ratings for Action to Meet Needs in Entry Sample (n = 153) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	$SE\ B$	β
Intercept	-3.99***	0.93	0.00
ABILITIES Index	0.08***	0.01	0.59
Male Gender	0.27	0.20	0.09
R^2		0.35	
F		41.08***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.2b **Regression of Male Gender Predicting COS Ratings for Action to Meet Needs in Entry Sample (n = 153) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	$SE\ B$	β
Intercept	-0.78	0.71	0.00
ABILITIES— Structural Integrity	0.10***	0.01	0.51
Male Gender	0.14	0.22	0.05
R^2		0.26	
F		26.68***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.3a   Regression of Race/Ethnicity Predicting COS Ratings for Action to Meet Needs in Entry Sample ($n = 153$) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-3.68***	0.94	0.00
ABILITIES Index	0.08***	0.01	0.58
Race/Ethnicity			
Hispanic	-0.38	0.29	-0.09
Black	-0.33	0.27	-0.08
Other	-0.08	0.46	-0.01
R^2		0.36	
F		20.57***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Exhibit 3.3b   Regression of Race/Ethnicity Predicting COS Ratings for Action to Meet Needs in Entry Sample ($n = 153$) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-0.66	0.71	0.00
ABILITIES— Structural Integrity	0.11***	0.01	0.52
Race/Ethnicity			
Hispanic	-0.34	0.31	-0.08
Black	-0.65*	0.29	-0.16
Other	-0.06	0.48	-0.01
R^2		0.29	
F		14.94***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Exhibit 3.4a   Regression of Child Age at Entry Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 153) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-4.26***	0.93	0.00
ABILITIES Index	0.08***	0.01	0.59
Child Age at Entry (months)	0.01	0.01	0.14
R²		0.37	
F		43.33***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.4b   Regression of Child Age at Entry Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 153) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-0.75	0.71	0.00
ABILITIES— Structural Integrity	0.10***	0.01	0.51
Child Age at Entry (months)	0.00	0.01	0.01
R²		0.26	
F		26.4***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.5a   **Regression of Family member presence at the COS Team Meeting Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 153) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-3.78***	0.92	0.00
ABILITIES Index	0.08***	0.01	0.59
Parent or family member at the COS Team Meeting	0.24	0.20	0.08
R^2		0.36	
F		40.20***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.5b   **Regression of Family member presence at the COS Team Meeting Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 153) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-0.69	0.71	0.00
ABILITIES— Structural Integrity	0.10***	0.01	0.50
Parent or family member at the COS Team Meeting	0.34	0.22	0.11
R^2		0.27	
F		26.12***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.6a   **Regression of Number of COS Team Members Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 153) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-3.79***	0.98	0.00
ABILITIES Index	0.08***	0.01	0.60
Number of COS Team Members	0.00	0.05	0.00
R^2		0.36	
F		39.15***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.6b   **Regression of Number of COS Team Members Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 153) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-0.60	0.78	0.00
ABILITIES— Structural Integrity	0.10***	0.01	0.51
Number of COS Team Members	-0.01	0.06	-0.01
R^2		0.26	
F		24.48***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.7a Regression of Service Coordinator on COS Team Predicting COS Ratings for Action to Meet Needs in Entry Sample (n = 153) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Action to Meet Needs		
	B	$SE\ B$	β
Intercept	-4.24***	0.99	0.00
ABILITIES Index	0.08***	0.01	0.60
Service Coordinator on COS Team	0.32	0.28	0.08
R^2		0.35	
F		40.71***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.7b Regression of Service Coordinator on COS Team Predicting COS Ratings for Action to Meet Needs in Entry Sample (n = 153) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Action to Meet Needs		
	B	$SE\ B$	β
Intercept	-1.38	0.81	0.00
ABILITIES— Structural Integrity	0.11***	0.01	0.53
Service Coordinator on COS Team	0.48	0.30	0.11
R^2		0.27	
F		28.13***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.8a   Regression of X on COS Team Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 153) (ABILITIES Index as Covariate)

Variable	B	SE B	β
	COS Entry Ratings: Action to Meet Needs		
Intercept	-3.96***	1.11	0.00
ABILITIES Index	0.08***	0.01	0.56
Early Intervention	0.35	0.40	0.11
Male Gender	0.20	0.21	0.06
Race/Ethnicity			
Hispanic	-0.28	0.30	-0.07
Black	-0.37	0.27	-0.10
Other	-0.10	0.46	-0.01
Child Age at Entry (months)	0.02	0.01	0.22
Parent at the COS Team Meeting	0.41	0.25	0.14
Number of COS Team Members	-0.07	0.07	-0.09
R^2		0.39	
F		10.00***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Exhibit 3.8b   Regression of X on COS Team Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 153) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-0.91	0.96	0.00
ABILITIES— Structural Integrity	0.09***	0.02	0.47
Early Intervention	0.52	0.43	0.17
Male Gender	0.12	0.23	0.04
Race/Ethnicity			
Hispanic	-0.24	0.32	-0.06
Black	-0.72*	0.29	-0.18
Other	0.02	0.49	0.00
Child Age at Entry (months)	0.01	0.01	0.16
Parent at the COS Team Meeting	0.57*	0.26	0.19
Number of COS Team Members	-0.07	0.07	-0.09
R^2		0.32	
F		6.94***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Note: Service coordinator participation was correlated significantly with EI program status and so was removed from the multiple regression model.

Appendix K

Levels of 7-Point COS Ratings with Mean Scores on BDI-2, Vineland-II, and ABILITIES Index at Program Entry in the Child Assessments Study

Data presented in this appendix extend information reported about results from Study 2. Mean scores on relevant subdomains of the BDI-2 and the Vineland-II, and for relevant subareas of the ABILITIES Index are provided for each of the 7 COS rating points. The sample size of 153 children was too small to create reliable estimates at every rating point, with estimates especially uncertain at either end of the distribution. So, these data should be viewed with caution. Data are provided for visual inspection and as background information for future research that might include larger samples of children. We expected to see a stair-step pattern such that higher COS ratings would be associated with higher assessment tool scores.

Positive Social Relationships

Across the three assessment tools, for positive social relationships, the majority of the 7 step pattern holds at each level (Exhibit K-1).

- For Vineland socialization, 5 out of 7 (71%) patterns hold. The exceptions are steady means rather than increases between 3 and 4 and a drop in means between 6 and 7.
- For BDI-2-social-emotional, 5 out of 7 (71%) patterns hold. The exceptions are higher than expected means for children rated as 1 and a drop in means between 6 and 7.
- For ABILITIES Index social/communication scores, 5 out of 7 (71%) patterns hold. The exceptions are higher than expected means for children rated as 1 and steady means rather than increases between 3 and 4.

Note that very few children received ratings of 1 so abberations at that step are not surprising.

Exhibit K-1   **Mean Assessment Tool Scores by 1-7 COS Entry Ratings on Positive Social Relationships in Entry Sample (*n* = 153)**
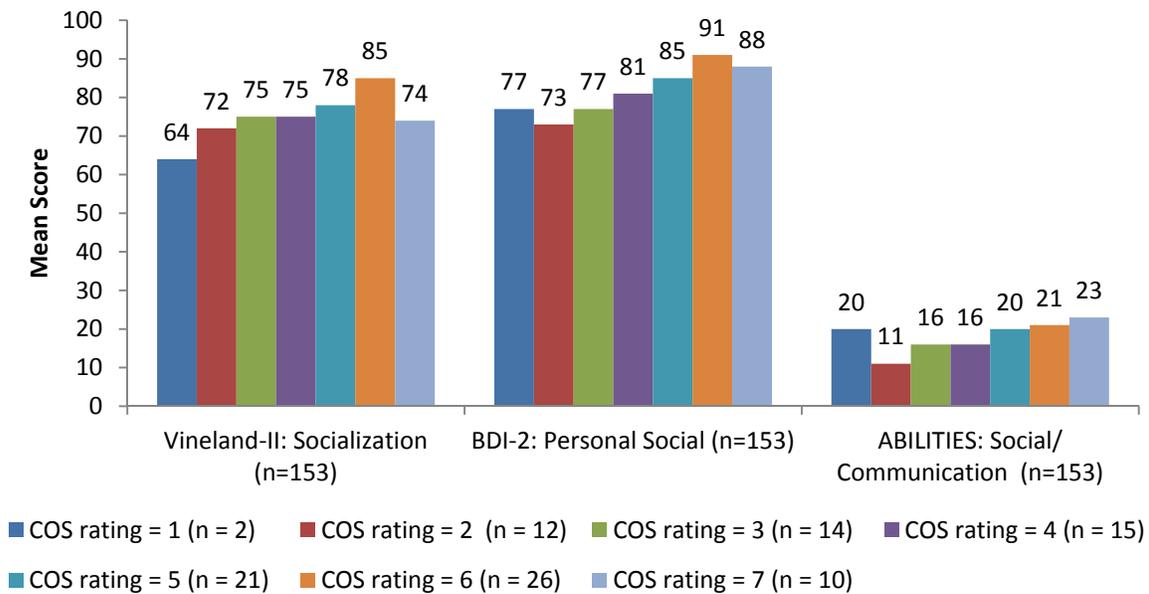
Knowledge and Skills

Across the assessment tools, for knowledge and skills, the majority of the 7-step pattern holds at each level (Exhibit K-2).

- For Vineland communication, 5 out of 7 (71%) patterns hold. The exceptions are higher than expected means for children rated as 1 and a drop in means between 5 and 6.

- For BDI-2 communication, 6 out of 7 (86%) patterns hold. The exception is higher than expected means for children rated as 1. Again very few children have COS ratings of 1, so that group is more likely to be unstable.

- For BDI-2 cognitive 4 out of 7 patterns hold (57%). The exceptions are higher than expected means for children rated as 1, higher than expected mean scores for five, and lower than expected mean scores for 7.

- For ABILITIES Index cognitive/communication scores, 5 out of 7 (71%) The exceptions are higher than expected means for children rated as 1 so that the score is the same as for a 2 and higher than expected mean scores for five, leading to the same mean score as for six.

Exhibit K-2   Mean Assessment Tool Scores by 1-7 COS Entry Ratings on Knowledge and Skills in Entry Sample (*n* = 153)
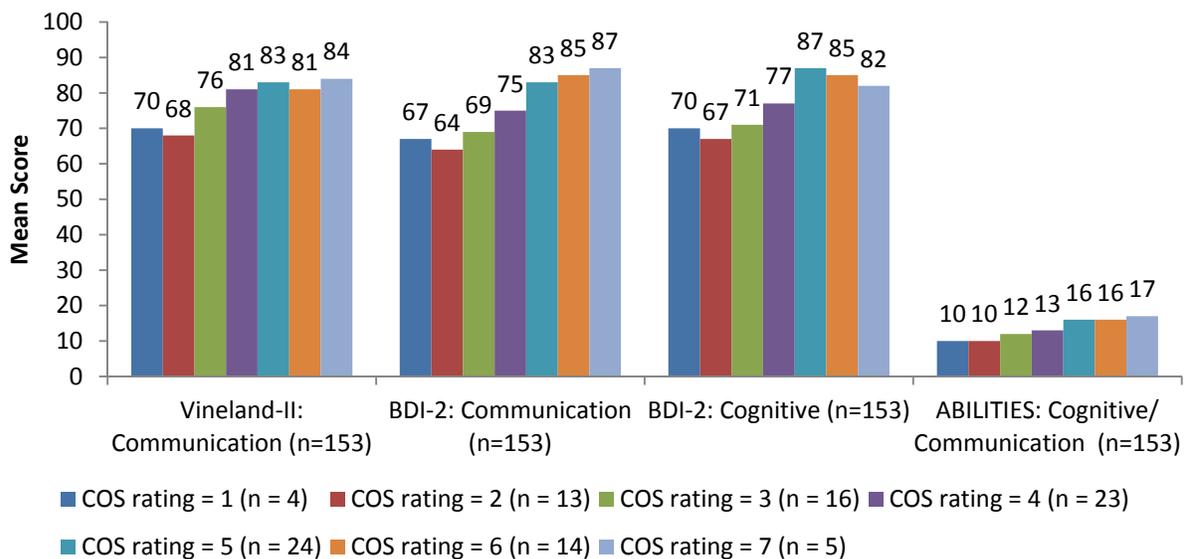
Action to Meet Needs

Across the assessment tools, for action to meet needs, the majority of the 7 step pattern holds at each level (Exhibit K-3).

- For Vineland daily living skills, 6 out of 7 (86%) patterns hold. The exception is higher than expected means for children rated as a 5, resulting in the same mean value as for a six.
- For BDI-2 adaptive, 5 out of 7 (71%) patterns hold. The exceptions are higher than expected means for children rated as 1 and lower than expected mean scores for the group of children with a 6.
- For ABILITIES Index structural integrity scores, 6 out of 7 (86%) The exception is higher than expected mean for children rated with 4 so that the score is the slightly above that for 5.

Exhibit K-3 **Mean Assessment Tool Scores by 1-7 COS Entry Ratings on Action to Meet Needs in Entry Sample (n = 153)**

Legend:
- COS rating = 1 (n = 3)
- COS rating = 2 (n = 12)
- COS rating = 3 (n = 12)
- COS rating = 4 (n = 22)
- COS rating = 5 (n = 28)
- COS rating = 6 (n = 15)
- COS rating = 7 (n = 8)
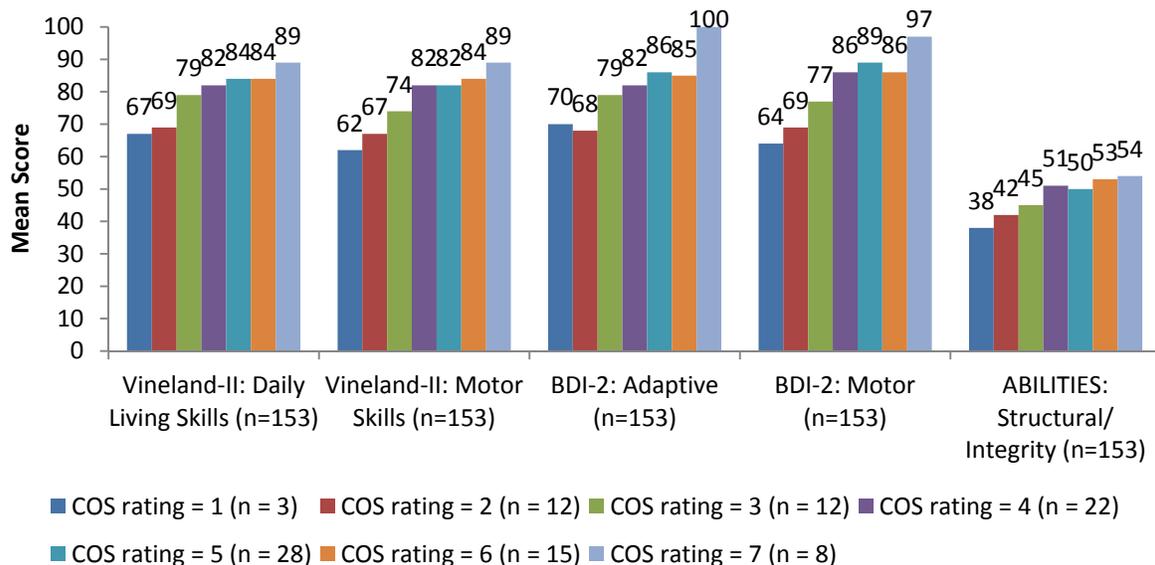
Summary

Evidence suggests that across the three assessment tools examined in relation to COS ratings on the outcomes, mean assessment tool scores for children with different COS ratings showed many of the expected patterns of increasing mean scores (i.e., a stair-step pattern) despite very small sample sizes on which to base mean estimates. Aberrations were most common at extreme rating points (especially for ratings of 1) where estimates were based on a smaller number of individuals. More research is needed with larger sample size to investigate relationships between assessment tools and COS ratings across the three outcomes and for EI and ECSE separately.

Appendix L

Team Decision-Making Study Project Developed Forms

ECSE Team Decision-Making Study: Meeting Information Form

Provider Information Form

Team Decision-Making Study Video Coding Form

ECSE Team Decision-Making Study
Meeting Information Form

Child name *(print)***:** *First* _____ *Last* _____

Date of birth: _____ / _____ / _____   **Date of COSF:** _____ / _____ / _____
_Mo_   _Day_   _Year_   _Mo_   _Day_   _Year_

Disability type: ☐Developmental Delay ☐Speech Language Impairment ☐Other: _____

Timing of COSF: ☐Entry ☐Exit

If exit, length of time child received services: ☐6-12 mos.☐>12 mos.

Type of meeting: ☐IEP ☐COSF only ☐Other (specify): _____

When was information relevant to the COSF discussed during the meeting?

☐Beginning   ☐Middle   ☐End   ☐Throughout

Person completing form:

First _____ *Last:* _____ *Email:* _____

Participants involved in meeting: *(In person, by phone, or some other way)*

Name *(Check box for service coordinator)*	**Role** *(e.g., parent, physical therapist, social worker)*	Study consent obtained?	Provider Info form obtained?
☐		☐	☐
☐		☐	☐
☐		☐	☐
☐		☐	☐
☐		☐	☐
☐		☐	☐
☐		☐	☐

Family information on child functioning *(check all that apply)***:**

☐Received in team meeting ☐Collected separately ☐Incorporated into assessment(s) ☐Not included

Please complete seating map on other side ➡

_____

(Office use) Child's ID: _____ Date received: _____ / _____ / _____

[Program name (ID), State, ECSE form]

Seating Map

Please sketch the table or the layout of the room where the meeting occurred to help us identify who is speaking in the video.

- *Note where the video camera was located.*
- *Write the name of each person in the location where he or she sat during the meeting.*
- *List the names of anyone who participated by phone.*

ENHANCE
Promoting quality child outcomes data

Provider Information Form

Program: | **ID:** | **State:**

Date completed *(mm/dd/yy):* ☐☐ / ☐☐ / ☐☐☐☐

Provider name *(first):* ☐☐☐☐☐☐☐☐☐ *(last):* ☐☐☐☐☐☐☐☐☐☐☐☐☐☐

Phone: ☐☐☐ / ☐☐☐ / ☐☐☐☐ *(Ext.)* ☐☐☐☐

Fax: ☐☐☐ / ☐☐☐ / ☐☐☐☐   **Email:**

Professional role(s): *(Mark (X) ALL your professional roles on IFSP/IEP teams.)*
- ☐ Early Interventionist/child development specialist/infant specialist   ☐ Occupational Therapist   ☐ Physical Therapist
- ☐ Psychologist   ☐ Service coordinator/family resource coordinator   ☐ Special Education teacher
- ☐ Speech-Language Pathologist   ☐ Social worker   ☐ Other (describe):

Preferred contact method *(for clarifying /follow-up questions only):* ☐ Phone   ☐ Email   ☐ Fax

---

1. **How long have you been providing services to young children birth to five with disabilities?**
   ☐ Less than 1 year   ☐ 1-2 years   ☐ 3-5 years   ☐ 6-10 years   ☐ 11 years or more

2. **Have you worked (in any capacity) with young children birth to five *without* disabilities?** *(e.g., child care, teaching, assessment):*
   ☐ No   ☐ Yes ▶ **If Yes, for how long?**   ☐ Less than 1 year   ☐ 1-2 years   ☐ 3-5 years   ☐ 6 years or more

Note: In some places the COSF is called the ECO form or the child outcomes questions.

3. **Approximately how many COSFs do you think you have participated in?** *(include as a team member or facilitating the team.)*
   ☐ Zero   ☐ 1-10   ☐ 11-30   ☐ 31-50   ☐ More than 50

4. **How many hours have you spent being trained on the COSF?**
   ☐ None   ☐ Less than 1 hour   ☐ 1-2 hours   ☐ 3-4 hours   ☐ 5-8 hours   ☐ 9-15 hours   ☐ More than 15 hours

5. **How well do you understand:** *(Mark (X) one for each item.)*

	Not at all well	Not well	Well	Very well
a. The skills and behaviors included in each of the three functional outcomes?	☐	☐	☐	☐
b. The definition of the 7 rating points?	☐	☐	☐	☐
c. The degree to which different skills and behaviors are age appropriate?	☐	☐	☐	☐

When completed: Give to local ENHANCE Program/District Liaison, or
FAX directly to 877-364-2620, or
Mail to ENHANCE Study, SRI International, 333 Ravenswood Ave. BS156, Menlo Park, CA 94025.

34061

(Office use) Provider ID: _____ Date received: _____ / _____ / _____

Team Decision-Making Study Video Coding Form

Child ID _____   Child First Name_____

Filename(s)_____   Child DOB __ - __ - __ Approx age: __ yr __ mo

Meeting Timing __ entry ___ exit __ dual (exit C & entry B)   Meeting Date __ - __ - __ Coding Date __ - __ - __

EI/ECSE Program ____   State ____ Program Name _____   Coder Name _____ Coding:_____

Video Format: ___ Team Discussion, including of child's functioning   ___Team discussion, not whole meeting on the tape
(select one)   ___ Individual explanation   ___ Other *(specify):*_____

Comments about team approach:

Documentation

If missing entire COS form, check here _____ and skip to next section.

1. Team rating on COS form   O1 ____   O2 _____   O3 _____
2. Progress   O1 ____   O2 _____   O3 _____
   (Only if exit or dual on meeting type)

	No	No, indicates evidence is in IFSP or IEP	Yes, minimally	Yes, more than minimally
3. Evidence is documented with rating on COS				
Outcome **1**				
Outcome **2**				
Outcome **3**				

If no on 3 a, b, & c then auto code as "No evidence on COS for three questions below" & skip presenting question 4a.b.c and 5 a, b, c.

	No evidence on COS	No	Yes
4. Evidence anywhere on COS *sufficient* to **justify rating**?			
Outcome **1**			
Outcome **2**			
Outcome **3**			

	No evidence on COS	No	Yes
5. Evidence anywhere on COS *consistent* with the **rating**?			
Outcome **1**			
Outcome **2**			
Outcome **3**			

6.Across the form, does **evidence** on the **COS documentation** categorize skills with **wrong outcomes**?	
Yes, more than minimal	
Yes, minimal	
No, no mistakes	
No, no evidence on form	

	No	Yes, but likely from earlier part of meeting not on tape	Yes	No Evidence on COS	Can't Tell
7. **More information on COS form than** discussed on **video** *(Inconsistency - expect more info in video than form. Inconsistent if there is info on form that was never in the discussion or evidence.) is assessment scores, but video never mention assessment tool, etc.) Includes new sources of info as well as specific content.*					
Outcome **1**					
Outcome **2**					
Outcome **3**					

8. Comments:

Meeting Length
1. Length of discussion/decision

Outcome #1 _____ minutes   Outcome #2 _____ minutes   Outcome #3 _____ minutes

   OR _____ minutes integrated across outcomes if outcomes discussion is too integrated to distinguish by outcomes

2. Key COS process taped   ___ hours ___minutes

3. Length of video/audio full meeting ___hours ___ minutes

Meeting Type
1. COS **Meeting Type**: ___ Embedded in IFSP/IEP   __ At End of IFSP/IEP
           ___ COSF only meeting ___ COS combo with some other meeting, can't tell mtg type
           ___ No meeting ____ at the end of evaluation meeting (if IFSP/IEP, COS, eval
      meeting combo, code as IFSP/IEP   Other: *(specify):*_____

1a. Comments:

2. Information from **parent** gathered by (*check all that apply*):
    __ participation in team meeting where COS decided *Based on video (if checked, complete next section)* (count even if parent is there, but says nothing)
    __ separate discussions (video/documentation)
     __through assessment tools (video/documentation)
     __ parent questionnaire of COS content (video/documentation)
    ___ other *(specify):* _____ (video/documentation)
    __ can't tell __ parent info not included in video
    (code based on info from documentation form not evident in video)

3. **Decision Approach**: __ Team discussion ___ Team discussion starting with COS form already completed by 2/more people
    (This question triggers skip patterns below) ___Asynchronous ___ Individual only decision __ Other *(specify):*_____

1. **Number of people present** in COS meeting (Video/Audio) _____ or ____ can't tell

   2. **Number of parents/guardians** present in COS meeting (Video/Audio): _____ or ____ can't tell

3. Based on all sources, **how many people contributed information to the COS process?** (Meeting information form list, documentation, or at the meeting, may share info that led to decision but not even participate in the meeting or the rating itself.) _____ or ____ can't tell

4. **Number of parents/guardians** who contributed information to the COS process. _____ or ____ can't tell

5. **Roles** of those **present in video** *(select all that apply from drop down menu)*: _____

6. Are those who contributed information **the same as those present in the video?**

____Yes, they are the same

____No, More contributed information

7. **Roles** of those who contributed information to the **COS process** (*Check all that apply from drop down menu):* _____

8. Was there a **service coordinator present**? __ no *(or at least nothing is checked on forms)* __ yes

9. Comments:

_____

When Parent is Included on Team
(If no parent participation in team meeting where COS decided, skip to next section)

1. Explain **why data** are being collected *(check all that apply)*:

__ No, not explained __Referenced earlier conversation about why __Yes, explain data are collected to observe child's progress ___Yes, explain data are collected for program improvement/accountability __ Yes, explain, but give other reason (If so, specify)_____ ____ Can't tell

2. Describes **meaning** of each **outcome** (*check all that apply*):
   __ No, not described ___ Referenced earlier description ___ Yes, for some, but not all, outcomes ___Yes, for each outcome __ Can't tell

3. Describes what **skills** are **expected** for a child **this age**: *(count if give list of skills and then say age level, even if not directly tied with words)*
   __ No, not described __Generically, not linkable to outcomes __ Yes, for some, but not all, outcomes ___Yes, for each outcome ___ Can't tell

4. Describes idea of **sequences** in development
   __ No __Yes __ Can't tell

5. **Parent spontaneously offers information** about child or rating during meeting
   __ No __ No, but can't confirm if nonverbal agreement/head nodding from video __ Yes, head nodding/agreement only
   __ Yes, spontaneously speaks up/shares ideas __ Can't tell

6. Providers **invite parent to comment/speak** during meeting (select one):
   __ No
   __ Yes, just ask if agree
   __ Yes, open-ended questions/statements, but not during COS discussion
   __ Yes, open-ended questions/statements encourage sharing information during COS part of meeting
   __ Can't tell

Quality of Team Process

(If individual only decision under decision approach, skip this section)

1. Overall, what was the quality of the general team process?
   lowest quality   1   2   3   4   5   highest quality
                          or   ___ can't tell

2. Comments:

3. Amount of parent input described or **shared** *(may be previously collected or parent participating in meeting, clear it represents parent perspective)*	__No __Yes, minimal __Yes, considerable __Can't tell				
4. How parent input was described or shared?	*(check all that apply)* __Neither ___ Professional shared/reported/read parent info __Parent spoke/participated ONLY in initial part of meeting, before COS was decided __ Parent spoke/participated __ Can't tell				
	Ratings				
5. What was the **extent of relevant contributions and dialogue** between team members (Professionals) ?	**1 -Low rating** Minimal or no dialogue or contributions between professionals	**2-Mid rating**	**3- High rating** Fully appropriate contribution from all professionals	**Only 1 professional/ no team**	**Can't tell**

	No	Yes	Appeared to agree based on form	Only 1 Professional	Can't Tell
6. Professionals on team **agreed on rating** *(e.g., seems to be genuine consensus, no signs someone did not agree at end)*					
Outcome **1**					

Outcome **2**					
Outcome **3**					

Quality of Child Outcomes Specific Process

1. What was the **quality of the Child Outcomes** specific **process**?	1 Lowest quality/ terrible example	2 Mid Rating	3 Moderate	4 Mid rating	5 Highest quality/ training tape	Can't Tell (not enough info to select #)
Outcome **1**						
Outcome **2**						
Outcome **3**						
Overall						

Consider individual components of the Child Outcomes process for each outcome. ONLY do overall if you can't do it by outcome… Or use overall only in situations where you can't really code each outcome individually

2. Comments:

3. **Considered Multiple settings /Situations appropriate for the child:**

3a. Settings/situations for Outcomes 1, 2, and 3: *(Check all that apply)*

	Outcome 1	Outcome 2	Outcome 3	Overall
Home (or parent report)	☐	☐	☐	☐
Assessment/eval (or professional report)	☐	☐	☐	☐
Preschool	☐	☐	☐	☐
Child care, including family day care situations	☐	☐	☐	☐
With other relatives (e.g., grandma's house, cousins, etc.)	☐	☐	☐	☐
With sibling(s)	☐	☐	☐	☐
Community (church, store, restaurant, etc.)	☐	☐	☐	☐
Community park, playground	☐	☐	☐	☐
Neighborhood situation (playing with friends who live next door, etc.)	☐	☐	☐	☐
With strangers (other than the assessment team)	☐	☐	☐	☐
Other setting--general, no specific location used	☐	☐	☐	☐
Other 1 (then specify for each of the "other" ratings) _____	☐	☐	☐	☐
Other 2 _____	☐	☐	☐	☐

Other 3 _____	☐	☐	☐	☐

3b. For **Outcome 1**, was consideration of settings/situations appropriate for the child?

    ___ No (not sufficient, including no probing to no possible settings)

    ___ Yes (mostly or completely)

    ___ Can't tell/Poor understanding of outcome, cannot rate (rate overall)

3c. For **Outcome 2**, was consideration of settings/situations appropriate for the child?

    ___ No (not sufficient, including no probing to no possible settings)

    ___ Yes (mostly or completely)

    ___ Can't tell/Poor understanding of outcome, cannot rate (rate overall)

3d. For **Outcome 3**, was consideration of settings/situations appropriate for the child?

    ___ No (not sufficient, including no probing to no possible settings)

    ___ Yes (mostly or completely)

    ___ Can't tell/Poor understanding of outcome, cannot rate (rate overall)

3e. Overall, was team consideration of settings/situations appropriate for the child?

    ___ No (not sufficient, including no probing to no possible settings)

    ___ Yes (mostly or completely)

    ___ Can't tell/Poor understanding of outcome, cannot rate (rate overall)

	Ratings			**Can't tell**	**Poor understanding of outcomes, can't rate for this outcome** (still rate overall)
	1 Not functional (only discrete skills or scores)	**2** Limited references of functional skills	**3** Mostly references functional skills		
4. Describe child's **functional** use of skills					
Outcome **1**					
Outcome **2**					
Outcome **3**					
Overall (use only if can't code by outcome)				■	

5. Description of **skills** which child is **not** yet **doing** (but would be expected or that come next)	___ No description ___ Yes, but only described on documentation ___ Yes, describe in video ___ Yes, describe on video and documentation ___ Can't tell

	Includes **major addition** of irrelevant skills	Includes **minor addition** or only relevant skills	No description of skills	Can't Tell
6. Description of outcome, functioning on outcome, and rating decision includes **skills relevant to that outcome area** *(major addition= has influence on the rating)*				
Outcome **1**				
Outcome **2**				
Outcome **3**				

7. Comments/which inappropriate skills:

	Can't Tell	No Description of Skills	Ratings		
			1 **Very limited breadth** *(focuses almost entirely on one aspect of outcome)*	**2** **Moderate Breadth** *(mixed, or missing one or more key areas)*	**3** **Good breadth** *(has key outcome areas, broad enough for good decision)*
8. Considered/discussed **breadth of outcome area** *(enough aspects of the outcome to be able to make a good decision about the child's functioning – based only on the video)* **Bolded** items = key areas					
Outcome **1** *(check all that apply* __relate with adults (parents/caregivers) __ relate with other adults (strangers, assessors, not primary caregivers) __relate with peers (6mo plus) __follow group rules/interacting with others (18 mo plus) __ social regulation of emotions/feelings __ language use in social situation ___ Other: _____					
Outcome **2** *(check all that apply* _ language __ memory/attention _ problem solving/reasoning (incl same/diff, matching) __understanding physical/social worlds (incl cause/effect, self understanding, community helpers, etc.) _early concepts/symbols _ preliteracy/preacademic skills _ general knowledge out of context __ other: _____					
Outcome **3** *(check all that apply* _ taking care of basic needs (showing hunger, dressing, feeding, toileting, etc.) _ contributing to own health/safety (24 mo plus) (e.g., follows rules, assists with hand washing, avoids inedible objects, car safety, playground etc) __getting from place to place (6 mo plus) (mobility, intentional movement to achieve aim) _ using tools (12 mo plus) (forks, pencils, strings, etc) __using language/sounds to indicate wants/needs __ other: _____					

9. Considered/discussed **enough depth** to have a good sense of the child skills in areas talked about	Ratings			No Description of skills	Can't tell
	1 No- Cursory or global discussion	**2** Yes mixed *One aspect described in depth, others brief*	**3** Yes, Appropriate *For any areas discussed – there is depth about skills*		
Outcome **1**					
Outcome **2**					
Outcome **3**					
Overall (code only if can't code by outcome)				██████	

10. Anchors skills using **rating criteria language** *(age appropriate, immediate foundational, foundational or like peers, younger child, much younger child)*	__No __Yes __No description of skills __Can't tell

11. Anchoring of skills **using sequences or ages** *(anchors, regardless of accuracy of anchors)*	__No __Yes __No description of skills __Can't tell

12. Professionals on team **inappropriately age-anchor skills** *(Major =did it wrong and it impacted rating)*	__Yes, Major __Yes, Minor __No inappropriate age anchoring __ **No age anchoring at all, regardless of whether or not appropriate** ___Can't tell

If yes, provide next questions (type of misunderstanding and influence) for completion. For no and can't tell, skip to the next question.
Types of misunderstanding with use of rating criteria observed:

13. Any indication that the professionals on the team do not understand the rating criteria?

	Yes	No	Can't tell due to limited process information
On each outcome was there any indication of misunderstanding or misapplication of rating criteria?			
Outcome **1**			
Outcome **2**			
Outcome **3**			

13a. Tell us about the following types of misunderstanding or misapplication of rating criteria observed.

	Outcome 1		Outcome 2		Outcome 3	
	Yes	**No**	**Yes**	**No**	**Yes**	**No**
Verbal reference suggests applying criteria the wrong way (e.g.), balances each other out						
Verbal reference overemphasizes one aspect of rating criteria too much (e.g., EVER)						
Application of rating criteria based on child's progress instead of distance from age-expected criteria						
Poor application of rating criteria due to problems with sequencing or anchoring age-level skills						
Poor application of rating criteria any other reason (e.g., may verbally describe criteria correctly but then rating doesn't match)						
Other misunderstanding (includes odd examples						

13b. In what ways did misunderstanding seem to interfere with a quality rating?

	Outcome 1		Outcome 2		Outcome 3	
	Yes	**No**	**Yes**	**No**	**Yes**	**No**
It led to categorizing child on the wrong side of the decision tree (1, 2/3, 4/5, 6/7)						
It led to confusion/error in decision between neighboring numbers						
Limited elicitation of examples or discussion						
It led to other influences						
	If Yes, specify_____		If Yes, specify_____		If Yes, specify_____	
No apparent influence on rating						

14. Decision tree explicitly used *(read, shown, in words, not just attached to COSF)*	__No __Yes, minimal reference _____ __ Yes, considerable use __Can't tell *(throughout all parts, consistent use, clear reliance on it)*

	No	Yes	Can't Tell
15. More than one rating was considered			
Outcome **1**			
Outcome **2**			
Outcome **3**			

	No	Yes
16a. Did one or more team members articulate a **rationale for the rating the team gave?** *(Note: if gave a rationale but not for the final rating option it does not count).*		
Outcome **1**		
Outcome **2**		
Outcome **3**		

For each outcome, if yes, go on to the next two questions. If no, skip further.

	No	Yes
16b. Was **Rationale** *sufficient* to justify the team's rating? (e.g., rationale gives enough information for why rating was given, provides enough information to discriminate significant of difference between neighboring scores; if can't tell, use no)		
Outcome **1**		
Outcome **2**		
Outcome **3**		

	No	Yes
16c. Was **rationale** *consistent with* the rating the team gave? (if consistent, use of rationale on the decision tree would lead to an option that included the rating given; inconsistent rationale of giving a child a 4 because she is almost ready to show age-expected behavior when it is expected she already has some for a rating of 4)		
Outcome **1**		
Outcome **2**		
Outcome **3**		

Comments in response to Q16, Q16b, and Q16c:

_____

_____

17. Any indication of an **explicit intent to alter ratings** (deflate entry ratings or inflate exit) *(based on comments and legitimacy in the rating direction – requires strong evidence)* ,	__Yes __No intent to alter ratings __Can't tell 17b. If yes, which, ___ deflate entry __ inflate exit __ other *(specify)* _____

18. Comments:

(Ask only if exit or dual meeting type) 19. At exit, child's **entry ratings** were described/**discussed**	__Yes __No entry ratings discussed __Can't tell 19b. . If yes, which outcome (check all that apply) __ can't tell __ O1 __ O2 __ O3

(Ask only if exit or dual meeting type) 20. At exit, **Child's progress** was discussed	__Progress not discussed __ Yes, for some but not all outcomes __ Yes, for each outcome __Can't tell

(Ask only if exit or dual meeting type) 21. At Exit, evidence of confusion with **"any progress"** idea	__Yes, evidence of confusion __ No, no evidence of confusion __Can't tell

22a. References **specific assessment tools**	__Neither __On documentation __ In video __ In video and documentation 22b. Names of tools: *(drop down menu)* _____ _____

23.Scores or age-levels on assessment tools mentioned	**No**	**Yes**
In Video		
On Documentation		

24. Specific **assessment tool content other than scores/age levels mentione**d in video	**1 –** **No specific Content**	**2 –** **Some specific content from assessments**	**3-** **Many instances of specific content from assessments**

(only code yes if very clear pulling it from tool, clear item not just behavior during assessment situation.)	

Consensus

Ask only if parent is part of team *(same question triggers as parent section)*	Yes	No unresolved disagreement	Can't Tell
1a. Active and **unresolved disagreement** on rating from one or more members of whole team *(team that includes a parent)*:			
Outcome **1**			
Outcome **2**			
Outcome **3**			

(If yes to any of 1a, 1b, 1c)	Yes	No
1b. Was unresolved disagreement **parent vs. professional**?		

2. Comments:

Ratings

Team Ratings *(1-7 options)*
1. Team rating in video   O1 _____ Can't tell__   O2 _____ Can't tell ___   O3 _____ Can't tell___

Coder range for ratings *(1-7 options) (Based on video, comment if additional information could consider on ABILITIES, COS, etc.)*

2a. O1 Single ___ RANGE: Low __ High __   Can't determine__   Coder confidence in amount/info available   low 1 2 3 high no rating
(Rating confidence based on amount/type of info)

2b. O2 Single ___ RANGE: Low __ High __   Can't determine __   Coder confidence in amount/info available   low 1 2 3 high no rating

2c. O3 Single ___ RANGE: Low __ High __   Can't determine __   Coder confidence in amount/info available   low 1 2 3 high no rating
available

2. Comments:

Video Characteristics/Future Reference:

1. Sound quality ___ good __ ok ___ poor

2a. Video quality ___ good ___ ok ____ poor ___ audio only

2b. Is everyone visible on the video? __no, miss a lot __ no, miss some __ yes

3a. Flag to consider when developing guidance document? __ yes __ no

3b. Comments:

4. Future training usefulness:
Useful example of **what <u>not</u> to do**   not at all useful

Useful example of **what is good to do**   not at all useful

5. Comments/noteworthy features (discussion, introducing topic, reaching consensus, etc.):

Additional comments/Coder notes:
1. About child/child characteristics:

2. About meeting/age-expectations/team rating:

3. Noteworthy for guidance documents:

4. Other notes:

5. Coding complete? __ yes __ no

Appendix M

Team Decision-Making Study Analyses: Relationships between Child and Team Characteristics and Entry COS Ratings on Each Outcome

Throughout this appendix we report findings from a series of regression analyses to investigate the extent to which specific child and team characteristics are related to COS ratings on each of the three outcomes at entry team decision-making study meetings ($n = 73$). These analyses consider the following outcomes: early intervention program, male gender, race/ethnicity, child age, family member presence, number of COS team members, and service coordinator, as well as a model showing these predictors in a combined model. Results from these analyses are summarized in the main report in Section 5 (Study 3).

- Each regression model was run the following two different ways:
  - With the ABILITIES Index total score (reversed) as a covariate, taking into account all aspects of the child's level of functioning
  - With the ABILITIES Index subarea[1] most relevant to the child outcome area as the covariate (i.e., Social/Communication for positive social relationships; Cognitive/Communication for knowledge and skills; and Structural Integrity for taking action to meet needs). Subarea scores also are reversed so that higher scores represent higher levels of functioning.
- Results are organized by the outcome area being predicted. Regressions are shown in pairs, with the ABILITIES Index total score as a covariate being reported first and displayed with an **orange** figure title and the regression with the ABILTIES Index subarea as a covariate reported next and a figure title in **blue**.

---

[1] More detail about the development of these ABILTIES Index subareas is contained in the main report.

1. Positive Social Relationships

Exhibit 1.1a   **Regression of Early Intervention Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 73) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	-3.44*	1.38	0.00
ABILITIES Index	0.08***	0.01	0.59
Early Intervention	-0.05	0.32	-0.01
R^2		0.35	
F		18.86***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.1b   **Regression of Early Intervention Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 73) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	0.40	0.53	0.00
ABILITIES Index - Social/Communication	0.26***	0.03	0.74
Early Intervention	-0.44	0.27	-0.13
R^2		0.54	
F		40.7	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.2a   Regression of Male Gender Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 73) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	**SE B**	**β**
Intercept	-3.45*	1.35	0.00
ABILITIES Index	0.08***	0.01	0.60
Male Gender	-0.28	0.35	-0.08
R^2		0.36	
F		19.34***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.2b   Regression of Male Gender Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 73) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	**SE B**	**β**
Intercept	0.38	0.55	0.00
ABILITIES Index - Social/Communication	0.25***	0.03	0.73
Male Gender	-0.23	0.30	-0.06
R^2		0.52	
F		38.54***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.3a   Regression of Race/Ethnicity Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 73) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	-3.00*	1.34	0.00
ABILITIES Index	0.08***	0.01	0.58
Race/Ethnicity			
Hispanic	-0.95*	0.44	-0.21
Black	-0.46	0.41	-0.11
Other	-0.59	0.52	-0.11
R^2		0.40	
F		11.22***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.3b   Regression of Race/Ethnicity Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 73) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	0.60	0.57	0.00
ABILITIES Index - Social/Communication	0.24	0.03	0.70
Race/Ethnicity			
Hispanic	-0.65	0.39	-0.15
Black	-0.23	0.36	-0.06
Other	-0.42	0.46	-0.08
R^2		0.54	
F		20.07***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.4a **Regression of Child Age at Entry Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 73) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Positive Social Relationships		
	B	$SE\ B$	β
Intercept	-3.39*	1.36	0.00
ABILITIES Index	0.08***	0.01	0.60
Child Age at Entry (months)	-0.01	0.01	-0.06
R^2		0.35	
F		19.17***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.4b **Regression of Child Age at Entry Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 73) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Positive Social Relationships		
	B	$SE\ B$	β
Intercept	-0.60	0.64	0.00
ABILITIES Index - Social/Communication	0.27***	0.03	0.77
Child Age at Entry (months)	0.02*	0.01	0.19
R^2		0.55	
F		43.51***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.5a   Regression of Family Member Presence at the COS Team Meeting Predicting COS Ratings for Positive Social Relationships in Entry Sample (*n* = 73) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	-3.56**	1.34	0.00
ABILITIES Index	0.08***	0.01	0.58
Parent or family member at the COS Team Meeting	0.40	0.32	0.12
R²		0.36	
F		20.06***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.5b   Regression of Family Member Presence at the COS Team Meeting Predicting COS Ratings for Positive Social Relationships in Entry Sample (*n* = 73) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	0.18	0.54	0.00
ABILITIES Index - Social/Communication	0.25***	0.03	0.71
Parent or family member at the COS Team Meeting	0.28	0.27	0.08
R²		0.53	
F		39.0***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.6a   Regression of Number of COS Team Members Predicting COS Ratings for Positive Social Relationships in Entry Sample (*n* = 73) (ABILITIES Index as Covariate)

	COS Entry Ratings: ositive Social Relationships		
Variable	*B*	*SE B*	*β*
Intercept	-3.38	1.37	0.00
ABILITIES Index	0.08	0.01	0.59
Number of COS Team Members	-0.03	0.07	-0.04
R²		0.35	
F		18.97***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.6b   Regression of Number of COS Team Members Predicting COS Ratings for Positive Social Relationships in Entry Sample (*n* = 73) (ABILITIES Index Subarea as Covariate)

	COS Entry Ratings: Positive Social Relationships		
Variable	*B*	*SE B*	*β*
Intercept	0.09	0.61	0.00
ABILITIES Index - Social/Communication	0.25***	0.03	0.73
Number of COS Team Members	0.04	0.06	0.05
R²		0.52	
F		38.29***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.7a　**Regression of Service Coordinator on COS Team Predicting COS Ratings for Positive Social Relationships in Entry Sample ($n = 73$) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	-3.48*	1.35	0.00
ABILITIES Index	0.08***	0.01	0.60
Service Coordinator on COS Team	-0.19	0.32	-0.06
R²		0.35	
F		19.1***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.7b　**Regression of Service Coordinator on COS Team Predicting COS Ratings for Positive Social Relationships in Entry Sample ($n = 73$) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	0.30	0.54	0.00
ABILITIES Index - Social/Communication	0.25***	0.03	0.72
Service Coordinator on COS Team	-0.11	0.28	-0.03
R²		0.52	
F		38.09***	

*p < .05, **p < .01, ***p < .001.

Exhibit 1.8a   Regression of Multiple Child and Team Characteristics Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 73) (ABILITIES Index as Covariate)

| | COS Entry Ratings: Positive Social Relationships | | |
Variable	*B*	*SE B*	*β*
Intercept	-1.67	1.52	0.00
ABILITIES Index	0.08***	0.01	0.59
Early Intervention	-0.77	0.55	-0.23
Male Gender	-0.15	0.35	-0.04
Race/Ethnicity			
Hispanic	-0.90	0.45	-0.20
Black	-0.76*	0.44	-0.19
Other	-0.63	0.54	-0.12
Child Age at Entry (months)	-0.03	0.02	-0.24
Parent at the COS Team Meeting	0.74	0.40	0.22
Number of COS Team Members	-0.14	0.09	-0.19
R^2		0.46	
F		5.84***	

*p < .05, **p < .01, ***p < .001.

Note: Service coordinator participation was correlated significantly with EI program status and so was removed from the multiple regression model.

Exhibit 1.8b   Regression of Multiple Child and Team Characteristics Predicting COS Ratings for Positive Social Relationships in Entry Sample (n = 73) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Positive Social Relationships		
	B	*SE B*	*β*
Intercept	-0.33	1.08	0.00
ABILITIES Index - Social/Communication	0.26***	0.03	0.75
Early Intervention	0.18	0.50	0.05
Male Gender	-0.27	0.31	-0.07
Race/Ethnicity			
Hispanic	-0.45	0.40	-0.10
Black	-0.25	0.39	-0.06
Other	-0.11	0.49	-0.02
Child Age at Entry (months)	0.02	0.02	0.24
Parent at the COS Team Meeting	0.34	0.36	0.10
Number of COS Team Members	-0.06	0.08	-0.08
R^2		0.58	
F		9.59***	

*p < .05, **p < .01, ***p < .001.

Note: Service coordinator participation was correlated significantly with EI program status and so was removed from the multiple regression model.

2. Knowledge and Skills

Exhibit 2.1a   **Regression of Early Intervention Predicting COS Ratings for Knowledge and Skills in Entry Sample (n = 73) (ABILITIES Index as Covariate)**

	COS Entry Ratings: Knowledge and Skills		
Variable	**B**	**SE B**	**β**
Intercept	-4.03***	1.23	0.00
ABILITIES Index	0.08***	0.01	0.64
Early Intervention	-0.42	0.29	-0.13
R^2		0.44	
F		27.43***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Exhibit 2.1b   **Regression of Early Intervention Predicting COS Ratings for Knowledge and Skills in Entry Sample (n = 73) (ABILITIES Index Subarea as Covariate)**

	COS Entry Ratings: Knowledge and Skills		
Variable	**B**	**SE B**	**β**
Intercept	0.95	0.47	0.00
ABILITIES— Cognitive/Communication	0.30***	0.03	0.71
Early Intervention	-0.79**	0.26	-0.25
R^2		0.53	
F		39.39***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Exhibit 2.2a **Regression of Male Gender Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 73) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	-4.38***	1.23	0.00
ABILITIES Index	0.08***	0.01	0.64
Male Gender	0.18	0.32	0.05
R²		0.43	
F		25.87***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.2b **Regression of Male Gender Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 73) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	0.62	0.51	0.00
ABILITIES—Cognitive/Communication	0.28***	0.04	0.68
Male Gender	0.12	0.31	0.04
R²		0.47	
F		31.14***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.3a   Regression of Race/Ethnicity Predicting COS Ratings for Knowledge and Skills in Entry Sample (n = 73) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	$SE\ B$	β
Intercept	-3.83**	1.22	0.00
ABILITIES Index	0.08***	0.01	0.63
Race/Ethnicity			
Hispanic	-0.64	0.40	-0.15
Black	-0.53	0.37	-0.13
Other	-0.93	0.47	-0.18
R^2		0.47	
F		15.02***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.3b   Regression of Race/Ethnicity Predicting COS Ratings for Knowledge and Skills in Entry Sample (n = 73) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	$SE\ B$	β
Intercept	1.01	0.53	0.00
ABILITIES— Cognitive/Communication	0.28***	0.04	0.66
Race/Ethnicity			
Hispanic	-0.57	0.39	-0.13
Black	-0.28	0.36	-0.07
Other	-0.54	0.47	-0.11
R^2		0.49	
F		16.43***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.4a **Regression of Child Age at Entry Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 73) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	-4.41***	1.24	0.00
ABILITIES Index	0.09***	0.01	0.65
Child Age at Entry (months)	0.00	0.01	0.03
R²		0.42	
F		25.68***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.4b **Regression of Child Age at Entry Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 73) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	-0.34	0.57	0.00
ABILITIES— Cognitive/Communication	0.31***	0.03	0.74
Child Age at Entry (months)	0.03**	0.01	0.25
R²		0.53	
F		39.64***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.5a   Regression of Family Member Presence at the COS Team Meeting Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 73) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	-4.39***	1.23	0.00
ABILITIES Index	0.09***	0.01	0.65
Parent or Family Member at the COS Team Meeting	0.12	0.29	0.04
R²		0.42	
F		25.75***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.5b   Regression of Family member presence at the COS Team Meeting Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 73) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	0.55	0.51	0.00
ABILITIES—Cognitive/Communication	0.28***	0.04	0.68
Parent or Family Member at the COS Team Meeting	0.25	0.28	0.08
R²		0.48	
F		31.74***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.6a **Regression of Number of COS Team Members Predicting COS Ratings for Knowledge and Skills in Entry Sample (n = 73) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Knowledge and Skills		
	B	$SE\ B$	β
Intercept	-4.12***	1.24	0.00
ABILITIES Index	0.09***	0.01	0.65
Number of COS Team Members	-0.07	0.07	-0.10
R^2		0.43	
F		26.7***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.6b **Regression of Number of COS Team Members Predicting COS Ratings for Knowledge and Skills in Entry Sample (n = 73) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Knowledge and Skills		
	B	$SE\ B$	β
Intercept	-4.12***	1.24	0.00
ABILITIES—Cognitive/Communication	0.09***	0.01	0.65
Number of COS Team Members	-0.07	0.07	-0.10
R^2		0.43	
F		26.7***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.7a **Regression of Service Coordinator on COS Team Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 73) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	-4.36***	1.18	0.00
ABILITIES Index	0.09***	0.01	0.67
Service Coordinator on COS Team	-0.69*	0.28	-0.21
R²		0.47	
F		30.77***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.7b **Regression of Service Coordinator on COS Team Predicting COS Ratings for Knowledge and Skills in Entry Sample (*n* = 73) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	0.88	0.49	0.00
ABILITIES—Cognitive/Communication	0.28***	0.04	0.68
Service Coordinator on COS Team	-0.46	0.28	-0.14
R²		0.49	
F		33.64***	

*p < .05, **p < .01, ***p < .001.

Exhibit 2.8a   Regression of Multiple Child and Team Characteristics Predicting COS Ratings for Knowledge and Skills in Entry Sample (n = 73) (ABILITIES Index as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	$SE\ B$	β
Intercept	-1.74	1.30	0.00
ABILITIES Index	0.08***	0.01	0.61
Early Intervention	-1.37**	0.48	-0.43
Male Gender	0.39	0.30	0.11
Race/Ethnicity			
Hispanic	-0.56	0.38	-0.13
Black	-0.88*	0.38	-0.22
Other	-1.20*	0.47	-0.23
Child Age at Entry (months)	-0.03	0.02	-0.29
Parent at the COS Team Meeting	0.62	0.35	0.19
Number of COS Team Members	-0.20*	0.08	-0.28
R^2		0.57	
F		9.22***	

*p < .05, **p < .01, ***p < .001.

Note: Service coordinator participation was correlated significantly with EI program status and so was removed from the multiple regression model.

Exhibit 2.8b   Regression of Multiple Child and Team Characteristics Predicting COS Ratings for Knowledge and Skills in Entry Sample (n = 73) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Knowledge and Skills		
	B	*SE B*	*β*
Intercept	1.13	0.98	0.00
ABILITIES—Cognitive/Communication	0.28***	0.04	0.68
Early Intervention	-0.64	0.48	-0.20
Male Gender	0.15	0.30	0.04
Race/Ethnicity			
Hispanic	-0.33	0.38	-0.08
Black	-0.49	0.38	-0.12
Other	-0.41	0.48	-0.08
Child Age at Entry (months)	0.01	0.02	0.12
Parent at the COS Team Meeting	0.63	0.34	0.20
Number of COS Team Members	-0.17*	0.08	-0.23
R^2		0.58	
F		9.73***	

*p < .05, **p < .01, ***p < .001.

Note: Service coordinator participation was correlated significantly with EI program status and so was removed from the multiple regression model.

3. Action to Meet Needs

Exhibit 3.1a **Regression of Early Intervention Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 73) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	**SE B**	**β**
Intercept	-3.64***	1.01	0.00
ABILITIES Index	0.08***	0.01	0.70
Early Intervention	-0.32	0.23	-0.11
R²		0.52	
F		37.32***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.1b **Regression of Early Intervention Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 73) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	**SE B**	**β**
Intercept	-0.42	1.04	0.00
ABILITIES— Structural Integrity	0.10***	0.02	0.51
Early Intervention	-0.26	0.29	-0.09
R²		0.29	
F		13.97***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.2a   Regression of Male Gender Predicting COS Ratings for Action to Meet Needs in Entry Sample (n = 73) (ABILITIES Index as Covariate)

	COS Entry Ratings: Action to Meet Needs		
Variable	*B*	*SE B*	*β*
Intercept	-3.88***	1.00	0.00
ABILITIES Index	0.08***	0.01	0.71
Male Gender	-0.11	0.26	-0.04
R^2		0.50	
F		35.59***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Exhibit 3.2b   Regression of Male Gender Predicting COS Ratings for Action to Meet Needs in Entry Sample (n = 73) (ABILITIES Index Subarea as Covariate)

	COS Entry Ratings: Action to Meet Needs		
Variable	*B*	*SE B*	*β*
Intercept	-0.69	1.01	0.00
ABILITIES— Structural Integrity	0.10***	0.02	0.52
Male Gender	0.04	0.31	0.01
R^2		0.28	
F		13.41***	

*$p < .05$, **$p < .01$, ***$p < .001$.

Exhibit 3.3a **Regression of Race/Ethnicity Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 73) (ABILITIES Index as Covariate)**

| Variable | COS Entry Ratings: Action to Meet Needs | | |
	B	*SE B*	*β*
Intercept	-3.56***	1.00	0.00
ABILITIES Index	0.08***	0.01	0.70
Race/Ethnicity			
Hispanic	-0.64	0.33	-0.17
Black	-0.26	0.30	-0.08
Other	-0.44	0.39	-0.10
R^2		0.53	
F		19.42***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.3b **Regression of Race/Ethnicity Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 73) (ABILITIES Index Subarea as Covariate)**

| Variable | COS Entry Ratings: Action to Meet Needs | | |
	B	*SE B*	*β*
Intercept	-0.40	1.00	0.00
ABILITIES— Structural Integrity	0.10***	0.02	0.53
Race/Ethnicity			
Hispanic	-0.77	0.40	-0.20
Black	-0.33	0.37	-0.09
Other	-0.63	0.47	-0.14
R^2		0.32	
F		8.16***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.4a **Regression of Child Age at Entry Predicting COS Ratings for Action to Meet Needs in Entry Sample (n = 73) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-3.98***	1.00	0.00
ABILITIES Index	0.08***	0.01	0.70
Child Age at Entry (months)	0.01	0.01	0.07
R²		0.51	
F		36.15***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.4b **Regression of Child Age at Entry Predicting COS Ratings for Action to Meet Needs in Entry Sample (n = 73) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-0.67	1.01	0.00
ABILITIES— Structural Integrity	0.10***	0.02	0.53
Child Age at Entry (months)	0.00	0.01	0.00
R²		0.28	
F		13.4***	

*p < .05. **p < .01 ***p< .001.

Exhibit 3.5a Regression of Family member presence at the COS Team Meeting Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 73) (ABILITIES Index as Covariate)

	COS Entry Ratings: Action to Meet Needs		
Variable	*B*	*SE B*	*β*
Intercept	-3.93***	1.00	0.00
ABILITIES Index	0.08***	0.01	0.70
Parent or family member at the COS Team Meeting	0.17	0.24	0.06
R²		0.51	
F		35.95***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.5b Regression of Family member presence at the COS Team Meeting Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 73) (ABILITIES Index Subarea as Covariate)

	COS Entry Ratings: Action to Meet Needs		
Variable	*B*	*SE B*	*β*
Intercept	-0.89	1.01	0.00
ABILITIES— Structural Integrity	0.10***	0.02	0.53
Parent or family member at the COS Team Meeting	0.35	0.28	0.13
R²		0.29	
F		14.47***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.6a **Regression of Number of COS Team Members Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 73) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-3.73***	1.01	0.00
ABILITIES Index	0.08***	0.01	0.71
Number of COS Team Members	-0.05	0.05	-0.08
R²		0.51	
F		36.29***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.6b **Regression of Number of COS Team Members Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 73) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-0.51	1.02	0.00
ABILITIES— Structural Integrity	0.10	0.02	0.53
Number of COS Team Members	-0.06	0.07	-0.09
R²		0.28	
F		13.93***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.7a **Regression of Service Coordinator on COS Team Predicting COS Ratings for Action to Meet Needs in Entry Sample (n = 73) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	$SE\ B$	$β$
Intercept	-3.89***	1.00	0.00
ABILITIES Index	0.08***	0.01	0.72
Service Coordinator on COS Team	-0.20	0.24	-0.07
R^2		0.51	
F		36.09***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.7b **Regression of Service Coordinator on COS Team Predicting COS Ratings for Action to Meet Needs in Entry Sample (n = 73) (ABILITIES Index Subarea as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	$SE\ B$	$β$
Intercept	-0.66	1.00	0.00
ABILITIES— Structural Integrity	0.10***	0.02	0.53
Service Coordinator on COS Team	-0.13	0.29	-0.05
R^2		0.28	
F		13.54***	

*p < .05, **p < .01, ***p < .001.

Exhibit 3.8a **Regression of Multiple Child and Team Characteristics Predicting COS Ratings for Action to Meet Needs in Entry Sample (*n* = 73) (ABILITIES Index as Covariate)**

Variable	COS Entry Ratings: Action to Meet Needs		
	B	*SE B*	*β*
Intercept	-2.58*	1.13	0.00
ABILITIES Index	0.08***	0.01	0.69
Early Intervention	-0.61	0.41	-0.22
Male Gender	-0.05	0.26	-0.02
Race/Ethnicity			
Hispanic	-0.54	0.33	-0.14
Black	-0.49	0.33	-0.14
Other	-0.44	0.40	-0.10
Child Age at Entry (months)	-0.01	0.01	-0.06
Parent at the COS Team Meeting	0.53	0.30	0.19
Number of COS Team Members	-0.16*	0.07	-0.25
R^2		0.58	
F		9.77***	

*p < .05, **p < .01, ***p < .001.

Note: Service coordinator participation was correlated significantly with EI program status and so was removed from the multiple regression model.

Exhibit 3.8b   Regression of Multiple Child and Team Characteristics Predicting COS Ratings for Action to Meet Needs in Entry Sample (n = 73) (ABILITIES Index Subarea as Covariate)

Variable	COS Entry Ratings: Action to Meet Needs		
	B	SE B	β
Intercept	0.64	1.09	0.00
ABILITIES— Structural Integrity	0.11***	0.02	0.58
Early Intervention	-1.07*	0.49	-0.38
Male Gender	0.15	0.30	0.05
Race/Ethnicity			
Hispanic	-0.69	0.39	-0.18
Black	-0.72	0.38	-0.21
Other	-0.74	0.47	-0.16
Child Age at Entry (months)	-0.03	0.02	-0.29
Parent at the COS Team Meeting	0.82*	0.35	0.29
Number of COS Team Members	-0.19*	0.08	-0.30
R^2		0.42	
F		5.14***	

*p < .05, **p < .01, ***p < .001.

Note: Service coordinator participation was correlated significantly with EI program status and so was removed from the multiple regression model.

Appendix N

ENHANCE Project Dissemination

Posters and Presentations Disseminated Via the ENHANCE Project

2014

Barton, L., Taylor, C., Spiker, D., Hebbeler, K., & Smyth, C. (2014, February). *Emerging evidence about the validity of the Child Outcomes Summary process.* Poster presented at the Conference on Research Innovations for Early Intervention (CRIEI), San Diego, CA.

Hebbeler, K., Barton, L., & Raber, S. (2014, July). *The power and challenges of early childhood integrated data systems (ECIDS): Implications for researchers.* Paper presented in symposium at Head Start's 12th National Research Conference on Early Childhood, Washington, DC.

Kahn, L., Hebbeler, K., Winer, A., Barton, L., & Spiker, D. (2014, July). *Longitudinal trends in data quality: An examination of trends in the quality of IDEA early childhood outcomes data.* Poster presented at Head Start's 12th National Research Conference on Early Childhood, Washington, DC.

Barton, L., Spiker, D., & Hebbeler, K. (2014, September). *Validity of the Child Outcomes Summary process: Updates from the ENHANCE project.* Presentation at the Improving Data, Improving Outcomes Conference, New Orleans, LA.

Smyth, C., Barton, L., Barrett-Zitkus, J., & Ashworth, J. (2014, October). *Team collaborations in discussions and decisions about children's functioning for accountability.* Poster presented at the Division for Early Childhood (DEC) Conference, St. Louis, MO.

2013

Barton, L., & Taylor, C. (2013, February). *Use of the Child Outcomes Summary process.* Presentation about ENHANCE findings at the Early Childhood Outcomes Center advisory group meeting. Washington, DC.

Hebbeler, K., Spiker, D., Barton, L., & Taylor, C. (2013, March). *Validity and accountability: Using survey data to examine the validity of early childhood outcomes measurement.* Presentation at the IES Principal Investigator Meeting, Washington, DC.

Barton, L., & Hodge, L. (2013, March). *Provider perceptions of the COS process.* Webinar for all participating ENHANCE programs and districts, Menlo Park, CA.

Barton, L., Taylor, C., Spiker, D., & Hebbeler, K. (2013, September). *Validity of the Child Outcomes Summary process: Updates from the ENHANCE project.* Presentation at the Improving Data, Improving Outcomes Conference, Washington, DC.

2012

Barton, L., Taylor, C., Hebbeler, K., & Spiker, D. (2012, February). *Validating Child Outcomes Summary ratings of young children's functional outcomes: Addressing methodological challenges.* Poster presented at the Conference on Research Innovations in Early Intervention, San Diego, CA.

Barton, L., & Taylor, C. (2012, February). *ENHANCE update: Research underway on the Child Outcomes Summary process.* Presentation at the Early Childhood Outcomes Center advisory group meeting, Arlington, VA.

Barton, L., Taylor, C., Hebbeler, K., & Spiker, D. (2012, February). *Comparing Child Outcomes Summary ratings to scores from assessment tools.* Poster presented at the Conference on Research Innovations in Early Intervention. San Diego, CA.

Barton, L., Backer, L., Cox, R., & Taylor, C. (2012, October). *Quality of child outcomes data: District experiences, state support, and national findings.* Presentation at the Division for Early Childhood Annual International Conference on Young Children with Special Needs and Their Families, Minneapolis, MN.

Taylor, C., Spiker, D., Barton, L., & Hebbeler, K. (2012, October). *Validity of state Child Outcomes Summary (COS) process ratings.* Poster presented at the Measuring and Improving Child and Family Outcomes Conference, Minneapolis, MN.

Barton, L. & Taylor, C. (2012, October). *Provider perceptions of the Child Outcomes Summary process.* Presentation at the Measuring and Improving Child and Family Outcomes Conference, Minneapolis, MN.

Smyth, C., & Barton, L. (2012, October). *Quality in the COS process.* Facilitated table top discussion at the Measuring and Improving Child and Family Outcomes Conference. Minneapolis, MN.

Barton, L., Taylor, C., Hebbeler, K., & Spiker, D. (2012, October). *Comparing Child Outcomes Summary ratings to scores from assessment tools.* Poster presented at the Measuring and Improving Child and Family Outcomes Conference, Minneapolis, MN.

2011

Spiker, D., Barton, L., Taylor, C., & Hebbeler, K. (2011, May). *Promoting quality child outcomes data*. Presentation at the International Society for Early Intervention, New York, NY.

Taylor, C., Barton, L., & Spiker, D. (2011, September). *Patterns in Child Outcomes Summary data: Analytic approaches and early findings from the ENHANCE project*. Paper presented at the Measuring and Improving Child and Family Outcomes Conference, New Orleans, LA.

2010

Barton, L., & Spiker, D. (2010, July). *Overview of ENHANCE*. Presentation at the Measuring Child and Family Outcomes Conference, Arlington, VA.