Child Outcomes Summary – Knowledge Check (COS-KC) Technical Report

December 2025

Authors:
Lauren Barton
Kathleen Hebbeler
Jenna Nguyen
Cornelia Taylor
Patrick Thornton

DaSy The Center for IDEA Early Childhood Data Systems

The contents of this report were developed under a grant from the U.S. Department of Education, #H373Z240001. The contents and resources do not necessarily represent the policy of the U.S. Department of Education, and you should not assume endorsement by the Federal Government. Project Officers: Meredith Miceli and Alexis Lessans.

December 2025

Acknowledgments:

We want to acknowledge the important role that the Early Childhood Technical Assistance Center (ECTA) played historically in early work that led to the development of the COS-KC. ECTA also provided helpful feedback throughout its development. Special thanks to the state leaders, and practitioners in those states who participated in the field test and/or pilot test of the COS-KC. These included people in Alabama, Alaska, Arkansas, California, Connecticut, Delaware, Hawaii, Idaho, Illinois, Maine, Maryland, Michigan, Minnesota, Mississippi, Nevada, New York, North Carolina, Ohio, Oklahoma, South Carolina, Texas, Vermont, Virginia, Washington, and Wisconsin. The generous sharing of time and feedback shared improved the tool for all who follow. Finally, we extend appreciation to Johns Hopkins University Center for Technology in Education for endless hours building the platform for the COS-KC field test and eventually the COS-KC itself.

Suggested citation:

Barton, L., Hebbeler, K., Nguyen, J., Taylor, C., & Thornton, P. (2025). *Child Outcomes Summary – Knowledge Check (COS-KC) technical report.* SRI.

SRI • 333 Ravenswood Avenue • Menlo Park, CA 94025-3493 • phone +1.650.859.2000 • www.sri.com

Contents

Introduction 1

History of the Development of the COS-KC 1

First Attempt: COS-CC 1

Change in Approach: COS-KC 2

COS-KC Item Development Process 2

Item Content 2

Quality Review of Item Content 2

Development of Different Forms for the Field Test 3

Field Test for Item Validation 4

Purpose 4

Methods 4

Instruments 4

Sample 5

Creation of the Prior COS Knowledge Index 5

Field-Test Results 6

Overall User Feedback on COS-KC Experience 6

Determining the Final Item Set for Each Form 7

Validity of the Final 30-Item Forms 8

Field Test Conclusion 11

Pilot Test for Usability of the COS-KC System 11

Purpose 11

Methods and Sample 12

Feedback and Improvements 12

Pilot Test Conclusion 12

Conclusion 13

References 14

Introduction

This report summarizes the development process and the testing that was undertaken to establish the validity of the Child Outcomes Summary – Knowledge Check (COS-KC). The COS-KC is a 30-item, multiple choice, pass-fail, open-book online assessment that examines whether a practitioner has sufficient knowledge to meaningfully participate in the COS process on a team to produce accurate ratings for child outcomes measurement. The assessment has six forms: three for practitioners working with newborns to 3-year-olds, and three for those working with 3-through 5-year-olds. An individual who does not pass on a first attempt is given a different form on their next attempt. The analyses contained in this report substantiate that the COS-KC is a valid assessment of an individual's knowledge of the COS process.

The assessment is part of the COS-KC system, which includes a set of reports that allow state and local administrators to track who has taken the assessment and how they performed. This report also summarizes the useability testing conducted with states to examine the features and utility of the COS-KC system. The results of the useability testing confirmed that the system functions well and that the reports in the system are easy to use and provide states and local programs with useful information.

History of the Development of the COS-KC

The COS-KC was developed in response to requests from the field for a way to understand if practitioners had knowledge of and were implementing the COS process as intended. Prior to developing the current COS-KC, the Center for IDEA Early Childhood Data Systems (DaSy) and the Early Childhood Technical Assistance Center (ECTA) tried to develop an assessment that would assess how well practitioners could apply their knowledge of the COS process in addition to how much they knew about the process. DaSy and ECTA developed and field-tested a COS-Competency Check (COS-CC) across several years before the data showed a much more comprehensive assessment was needed to measure application of knowledge. The centers then decided to use some of the items from the COS-CC and to assess only knowledge rather than competency, leading to the development and field test of the COS-KC.

First Attempt: COS-CC

The initial assessment, the COS-CC, was developed and field-tested from 2014 to 2018. This assessment consisted of two levels: a Level 1 foundational knowledge check with multiple choice questions, and a Level 2 activity in which individuals read a child example and applied their COS knowledge by answering questions about the evidence in the example and assigning an appropriate rating. The assessment provided information about the child's functioning from multiple providers and family members.

The field test showed that the assessment could not fully re-create the true COS process because the individual practitioner taking the assessment ultimately had to assign the child a rating without the benefit of asking questions of, having discussion with, or reaching consensus with other team members. The general conclusion from the field test was that the Level 1 questions passed basic validation standards, but the Level 2 questions did not perform well enough to be considered valid. Performance on Level 2 varied based on the example child's age and details, the background of practitioners, and the age-anchoring resources referenced. Participants also shared feedback that the assessment was time-consuming, it was not an

adequate representation of the actual COS process, and it was much harder than the actual team-based COS process with a child whom team members had observed and knew.

Based on these results, DaSy and ECTA decided to drop Level 2 of the assessment and pursue developing an assessment modeled on Level 1.

Change in Approach: COS-KC

The newly revised assessment was named the COS-KC, reflecting the focus on knowledge. The following sections describe the development process for the COS-KC items and the field-test validation of the items.

COS-KC Item Development Process

Item Content

The goal of the COS-KC is to identify whether a practitioner has the overall knowledge needed to understand and meaningfully participate in the COS team's decision-making process and produce accurate ratings. A team of technical assistance providers involved in creating and delivering professional development on the COS process discussed and decided on the content to be included in the COS-KC. The team reviewed content included in the national COS online module, in other COS professional development resources, and in Level 1 of the COS-CC. The team decided that the following topics were essential to meaningful participation in the COS process:

- Global content about why outcomes data are collected and why the COS process is needed
- The three child outcomes, functionality of skills, and how skills are related to the three outcomes
- Expectations for COS teaming, including the teaming and rating discussion process
- The 7-point COS rating scale and the definitions, or criteria, for each of the points on the rating scale
- Special considerations in assigning appropriate ratings on the 7-point scale, including what should and should not be considered in rating decisions
- Age anchoring, including what is meant by age-expected (AE), immediate foundational (IF), and foundational (F) skills
- Providing evidence or documentation for ratings and why it is important

For each topic, the team developed multiple choice items that included distractor responses representing common misunderstandings of the concepts or alternative concepts that are sometimes confused. All questions were written in English.

Quality Review of Item Content

Technical assistance providers with COS knowledge who had not been involved in the initial development process reviewed the items to ensure each item reflected the following features:

- A single correct or "best" answer
- Readability with conciseness and appropriate grammar

- Positive phrasing that minimizes the use of negatives or “not” in the questions
- Content that is clearly connected to COS process knowledge
- Clarity of the content being assessed—that is, questions are straightforward without a sense of being “trick” questions
- Only one concept addressed within each question
- Independence of items such that one item or decoy does not provide clues about another item on the assessment
- Plausibility of distractor options that are clearly wrong to those who understand the content but seem plausible to those less familiar with the content
- Universality of the concept and language used across states with variations in COS implementation processes
- Response options that are similar in form, length, and construction such that the correct answer does not stand out from the others
- Response options that are presented in a logical order, with the correct answer regularly appearing across all four response options, and
- Any content describing child functioning matched the age group of the test (e.g., examples represent children birth to 3 years old or 3 through 5 years old, consistent with the version of the assessment).

Items that raised concerns were revised or discarded. The item bank was then shared with five other COS professional development experts for review to identify items with issues and help select the best items to include in the field test.

Development of Different Forms for the Field Test

The item development team constructed six COS-KC forms of 42 items each for use in the field test. The team designed the field test to provide information about more items than needed for the final version of the COS-KC so that the best items could be selected. Three of the forms were for participants working with children birth to 3 years old, and three were for participants working with children 3 through 5 years old. Much of the content was the same, but if questions included an example of a child engaging in an activity, then the development team identified the approximate age level of the activity and the item was only used on the relevant form (birth to 3 or 3 through 5).

To maximize equivalence across versions, the field-test forms included a similar number of items about each of the content topics. Specifically, each form had 3–4 items on the global content, 8 items about the three child outcomes and functionality of skills, 4–5 items about COS teaming, 14–15 items about the 7-point scale and its rating criteria, 4–5 items on special considerations when determining ratings, 5–6 items on age anchoring, and 1 item on evidence and documentation. A higher number of items for a given topic reflected the team’s judgement of the importance of that topic to participation in the COS process.

The six forms of the COS-KC for the field test were provided to Johns Hopkins University Center for Technology in Education, a subcontractor to SRI, who built the technical system used for the consent process, registration form, assessment, and the user survey feedback form for the field test. This technical system was subsequently revised and serves as the basis for the current COS-KC.

Field Test for Item Validation

Purpose

The purpose of the field test was to evaluate, validate, and improve the COS-KC assessment. The field test examined the psychometric properties of the items and gathered participants' feedback on their experience taking the assessment.

Objectives for the field test were:

- To identify a final set of psychometrically sound items as evidenced by:
  - Items that had an appropriate range and level of difficulty, such that neither too many nor too few people scored any individual item correctly;
  - Items that distinguished between participants with higher and lower levels of prior COS knowledge; and
  - Items that had minimal comments about alternative interpretations or confusing wording in comments from those taking the assessment during the field test.
- To establish a cut point for passing based on the data collected.
- To test the following validity claims about the final set of items and the cut point for passing:
  - Each of the six forms is psychometrically sound:
    - Each form has an appropriate level of internal consistency—that is, a correlation among the items consistent with a set of items that represent the overarching construct of COS knowledge.
    - Each form contains items that span a range of difficulty, with individuals with the highest scores more likely to get the more difficult items correct.
  - Individuals with higher levels of prior COS knowledge pass the assessment at higher rates than those with lower levels.
  - The three forms for birth to 3 years old and the three forms for 3 through 5 years old are equivalent, as evidenced by the form not being a significant predictor of passing the assessment.

Methods

Instruments

The field test was conducted via an online platform where users created an account and signed in, reviewed and provided informed consent, answered a series of registration questions about their backgrounds, answered 42 multiple choice questions about the COS process, and completed a brief feedback form about their experience taking the assessment. The 42-item sets were expected to contain enough high-quality items from which to identify a set of 30 items for each of the final COS-KC forms.

Registration Questions

The registration questions asked about the user's state, role, length of time working with children with and without disabilities, primary ages of children they work with, preferred form of

assessment to take (birth to 3 or 3 through 5), whether English was their primary language, number of hours of prior COS training, type of training received, experience with the national COS online module, involvement in training on the COS, and number of COS ratings they had been involved in. These questions remain on the current COS-KC registration for new users.

The COS-KC Assessment

Users answered 42 multiple-choice items. Users were randomly assigned to receive one of three forms for either birth to 3-year-olds or 3 through 5-year-olds, based on their response to the registration question about the primary ages of children they work with.

Feedback Form

The feedback form asked users about the internet browser they used, any technical problems encountered during the field test, and experiences with saving and returning to the assessment. Users indicated how much they agreed or disagreed with statements about the clarity of the instructions, the helpfulness of the progress bar, and the ease of navigating through the COS-KC. Other questions asked about their use of and potential improvements to the resources provided during the COS-KC, as well as the amount of time it took to complete the assessment. Finally, there were opportunities for users to provide open-ended information about general reactions, comments or questions; questions or terms that were confusing; the flow or sequence of the assessment; and anything else that was important to know.

The instruments and methods were approved by an institutional review board (IRB) for research involving human subjects at SRI and at Johns Hopkins University.

Sample

Recruitment for participation in the COS-KC field test began in February 2021. The team sent a flyer and Q&A document outlining the plans for the field test and inviting participation to early intervention and early childhood special education leaders in states using the COS. Information about the field test also was shared at a COS Learning Community meeting. Interested state leaders nominated programs or individuals for participation and/or sent emails encouraging individuals or programs/districts to participate.

Individuals who took the assessment received a $20 gift card along with the opportunity to enter a raffle for a $100 gift card, which was given to six randomly selected individuals. Participants had the option to be in the raffle without completing the assessment. In states where individuals were not allowed to accept gift cards, alternative arrangements were made, such as donating books to a program in a similar dollar amount.

The final sample included 352 participants who completed at least some part of the assessment, with 187 participants from 17 states taking the forms for birth to 3-year-olds and 165 from 14 states taking the forms for 3- through 5-year-olds. A total of 296 participants completed the entire assessment and all of the registration items (158 across the birth to 3 forms, and 138 across the 3 through 5 forms).

Creation of the Prior COS Knowledge Index

The variable Prior COS Knowledge Index variable was created to capture prior training and experience related to the COS process, which was used as a proxy for prior COS knowledge. The team expected that higher Prior COS Knowledge Index scores would predict better

performance on the assessment. Participants were categorized as high, medium, or low on the index based on what they reported about their background experience with the COS on the registration form.

Table 1. Criteria for Categorizing Prior COS Knowledge of Participants

Prior COS Knowledge Index Category	Criteria
High	• Reported "9 or more total hours" of COS training regardless of their role as a COS trainer, OR • Reported "5–8 total hours" of COS training AND reported either: – "Conducting group trainings" or "developing materials to train others on the COS," OR – That they "are responsible for COS training with an entire region or state."
Medium	• Reported "3–4 total hours" of COS training AND reported they had completed "all" of the national COS online module; OR • Reported "5–8 total hours" of COS training AND reported they: – Are NOT either "conducting group trainings" or "developing materials to train others on the COS," AND – Are NOT "responsible for COS training with an entire region or state"
Low	• Reported "none," "less than 1 hour," or "1–2 total hours" of COS training, regardless of their answers to the other questions; OR • Reported "3–4 total hours" of COS training AND, when asked about the national COS online module, reported "yes, completed some of it," "no, have not completed it," or "unsure."

Application of these criteria resulted in 112 participants (32%) categorized as "high," 79 (22%) categorized as "medium," and 140 (40%) categorized as "low." Twenty-one participants (6%) were missing data needed to determine a category.

Field-Test Results

Overall User Feedback on COS-KC Experience

Participants' feedback from the COS-KC field test was mostly positive. On the Feedback Form, participants indicated that the COS-KC was a helpful refresher that covered all the core content needed for the COS process and that it "made them think." They liked the flow, finding it well organized and easy to follow. A number of participants commented that it was important for the COS-KC assessment to continue to be "open book." Many participants found the COS Process Quick Reference Guide to be a helpful resource for the assessment and indicated they would continue to use it afterward as a reminder of key content.

Several participants commented that some questions were long and wordy or confusing, with some pointing to specific items. On average, participants took about an hour to complete the 42 items, and some reported that the assessment took a long time to complete. Others expressed concerns that what was described in the assessment did not reflect how the COS process was implemented in their state. Examples differences included decisions about including families in the team decision-making process and state-specific language for terms that would be helpful to know about in advance. A few participants offered feedback on the

technical features of taking the assessment and improvements for better online functionality and movement. Although scores were not computed and participants did not receive notice if they had "passed" or "not yet passed" during the field test, many were eager to know their score.

Determining the Final Item Set for Each Form

The process of determining the final items for each of the COS-KC forms involved multiple steps and used multiple pieces of information from the field test. To identify the final item set, the team carried out the following activities:

- Generated a spreadsheet showing all of the items on which to record information to flag items for further review.
- Looked at field-test information for items on each form that were outliers with an especially high or low percentage of participants getting the answer correct or skipping the item altogether. Outliers were flagged.
- Compared the proportion of participants who were "high" on the Prior COS Knowledge Index and got each item correct, with the proportion who were "medium" or "low" on the index and got the item correct. The team expected to see some differentiation in correct responses among participants with higher prior COS knowledge. Items for which there was not a clear distinction were flagged.
- Reviewed correlations between getting an item correct and total score on the form. The team expected to see higher correlations between getting each item correct and getting a higher total score on the form. Items with low correlations between the item and the total score were flagged.
- Conducted item response theory analyses. The team reviewed the Wright Map for each form carefully to examine where each item fell in terms of difficulty. Items that had extremely high difficulty or extremely low difficulty were flagged.
- Looked at the Feedback Form data for comments from participants about specific items or item characteristics. The team noted these items and recorded details about them.
- Reviewed general comments from the Feedback Form and coded information about concerns where no item number was provided. For instance, several participants said it was hard to answer questions where "not" was included in the question stem. The team coded all such questions for careful review.
- Examined all the questions on each content topic together to make decisions about items within each topic area. The team identified a target range for the number of questions on each content topic for inclusion in the final 30-item set. Topics considered most essential to COS participation and determining accurate ratings were assigned a higher number of items.

Using the information compiled from the activities described above, two team members independently worked through the spreadsheet item by item and made recommendations to "keep item as is," "drop the item," or "modify the item." Items kept were those without flags that worked well or those on topics known to be difficult for participants to master during professional development but essential to applying rating criteria correctly. Items dropped were those that were flagged because they could be interpreted in a way that more than one answer could apply or those that were less clear than alternate items on a similar subject.

The team also reviewed the COS Process Quick Reference Guide and COS-KC: Key Terms, both of which are provided as resource links within the COS-KC and on the COS-KC: For Practitioners – Resources page. The team checked these resources to be sure that the wording

was consistent with the COS-KC items and correct answers and that the terms that participants said were less familiar in their states were fully defined. As part of this review, the team added information into these resources on any topics that were not previously addressed but were on the assessment.

The two team members met to discuss their recommendations and, in rare cases of disagreement, to reach consensus on recommended actions. Modifications included, for example, fixing a typographical error, reordering the response options so users read distractors before the actual answer, reordering words to remove "not" from sentences, and replacing a distractor response that users found confusing with a different wrong answer from an alternative item covering similar content. The team members reviewed the near-final list again to check if changes resulted in the target of 30 items on each form and a similar balance of content included across the forms. This first round of changes resulted in dropping 11 items from Forms 1 and 2, 7 items from Form 3, 11 items from Forms 4 and 6, and 15 items from Form 5. The team members conducted a second review similar to the first to be sure decisions would balance dropping items from various forms and result in a similar number of questions across each form. Consensus was reached to drop one more item each from Forms 1, 2, 4, and 6; to drop five more items from Form 3; to leave in one of the initially dropped items in Form 5; and to add two high-functioning items from other forms to Form 5. This resulted in a consensus list of 30 items for each of the six forms.

After two team members had a consensus list of items, they brought the information and recommendations back to the initial item development team to discuss decisions collectively. That group also reached consensus. This group also reviewed the COS Process Quick Reference Guide again to ensure the content in that resource was not misleading for any of the COS-KC questions.

The outcome of this process was six forms with 30 items. Five of the six forms contained a subset of the 42 items that were used in the field test for those forms. The team was only able to use 28 items from Form 5; the final two items added to that form were high-functioning items tested on a different form for the same age group.

Validity of the Final 30-Item Forms

After identifying the recommended 30 items for each of the six forms, the team set a pass rate and re-ran the analysis to determine the validity of the revised assessment.

Determination of Pass Rate for COS-KC

The team began the process of determining an appropriate pass rate for the COS-KC by asking how many items a user could miss and still convey that they had sufficient knowledge to appropriately participate in the COS process and determine accurate ratings. The team also reviewed how many items related to different kinds of content were within the assessment. Consensus was that missing five or fewer items—or a pass rate of 83%—would indicate the user did have enough knowledge to actively participate in the COS process and determine an accurate rating.

The team also analyzed data to see what pass rate would yield the most meaningful results. Using the 42-item dataset with each form, a simple kappa was computed between whether or not participants passed the COS-KC at 1 percentage point intervals from 75% to 100% and if they were high on the Prior COS Knowledge Index. The greatest correspondence occurred when setting the pass rate at 81%–83%. The kappa at this threshold was .3157 (n = 316 with complete data) which was indicative of fair-to-moderate correspondence between passing the

assessment and having high prior knowledge. Given the alignment between the results of the data analysis and the team's determination of the appropriate passing rate, 83% was applied in subsequent analyses as the pass rate.

Statistical Analyses

Internal Consistency. Because each form of the COS-KC was developed to represent one general overarching construct of knowledge of the COS process, the items were expected to be broadly correlated with one another and to function similarly. Overall raw Cronbach's alpha values ranged between .67 and .72, as follows:

- Birth to 3-year-olds: Form 1 ᾱ (*n* = 30) = .72; Form 2 ᾱ (*n* = 30) = .68; Form 3 ᾱ (*n* = 30) = .67.
- 3- through 5-year-olds: Form 4 ᾱ (*n* = 30) = .71; Form 5 ᾱ (*n* = 28) = .70; Form 6 ᾱ (*n* = 30) = .69.

These values hover right around the acceptable range, with widespread agreement that .70 is acceptable and some research referring to .65 as minimally acceptable or even .60 as an acceptable index (Konting, 2009; Nunnally & Bernstein, 1994; Pallant, 2001). Given that the team was looking to develop a short assessment and that the alpha values reflect the number of items, the team concluded that the alphas were acceptable.

Equivalence of Forms. Rasch analyses provided consistent evidence that the final COS-KC item sets function well across all six forms. Across forms, infit mean-square values for retained items fell within 0.75–1.44, well inside accepted fit thresholds for productive measurement and with no items approaching the >2.0 misfit boundary (Linacre, 2002). Item difficulty values across the six forms collectively ranged from –3.4 to +2.3 logits, providing broad coverage of the COS-KC ability continuum (Wright & Stone, 1979). Item reliability values across forms fell between .75 and .82, indicating stable item calibration (Wilson, 2004), and person reliability values ranged from .42 to .70, consistent with expectations for short dichotomous knowledge assessments (Hambleton et al., 1991).

Exact-match rates between item difficulties and people's abilities across forms were high, generally between 80% and 96%; Wright Maps also showed close alignment between person and item means on all forms, confirming strong targeting (Wright & Masters, 1982).

Pass Rates by Form. Using the agreed-upon pass rate, the team analyzed field test data to determine if the percentages of participants who passed the COS-KC using various forms were similar. Table 2 shows the number and percentage of participants who passed the 30-item COS-KC by form. The range across forms was 61% to 69% for the 0–3 forms and 62% to 63% for the 3–5 forms of the assessment.

Based on these results showing that the pass rates fall within a very similar band, there is no indication that the 0 to 3 forms (Forms 1–3) differ meaningfully from the 3–5 forms (Forms 4–6). This pattern supports the conclusion that the forms are performing equivalently across age groups. Looking at the item difficulty distributions supports the same conclusion as the pass rates: The forms do not differ meaningfully between the 0 to 3 versions and the 3–5 versions. All six forms show highly overlapping item difficulty ranges (approximately –3.5 to +2.5 logits across forms), with no form or population set showing systematically easier or harder items. This pattern is consistent with the goal of producing parallel forms and supports the conclusion that the forms function equivalently across age groups.

Table 2. Percentage Passing the 30-Item COS-KC by Form

Form	Population	*n*	Passed (Number)	Passed (Percent)
Form 1	EI providers (0 to 3)	62	38	61
Form 2	EI providers (0 to 3)	41	27	66
Form 3	EI providers (0 to 3)	55	38	69
Form 4	ECSE providers (3 - 5)	46	29	63
Form 5	ECSE providers (3 - 5)	47	29	62
Form 6	ECSE providers (3 - 5)	45	28	62

Note. Includes only those with complete data on the final dataset. EI = Early Intervention ECSE = Early Childhood Special Education.

Predictors of Passing or Not Passing the COS-KC. In the initial 42-item field-test version, participants with higher self-reported prior COS knowledge performed substantially better than those with lower knowledge levels. Pass rates were 28% for low-knowledge participants, 37% for medium-knowledge participants, and 64% for high-knowledge participants. Although these values come from the full field-test item set rather than the final calibrated forms, they provide early evidence that the assessment appropriately differentiates users based on external indicators of COS expertise.

Logistic regression analyses investigated several key factors to determine if they were related to whether or not individuals passed the final 30-item COS-KC. Analyses were conducted separately for data for the 0–3 forms and the 3–5 forms. For the 0–3 forms, the full model, including the forms and information about prior COS knowledge, fit the data significantly better than the reduced model (LRT χ^2= 9.282*, *df* = 3, $p < .05$; Wald χ^2 = 8.227* = $p < .05$). The form used was not a significant predictor of passing the assessment. However, having high prior COS knowledge was a significant predictor relative to people who were categorized as having low or medium prior knowledge (Table 3).

Table 3. Logistic Regression Results for 0 to 3 Forms of the COS-KC Assessment (*n* = 158)

Source	*df*	B Estimate	*SE* B	Wald χ^2	*p*	Odds Ratio (OR) Estimate	95% CI OR Minimum	95% CI OR Maximum
Intercept	1	0.2898	0.3573	0.6575	.4174	–	–	–
Form 1	1	-0.1252	0.4329	0.0836	.7725	Form 1 vs. 2: 0.882	0.373	2.055
Form 3	1	0.2557	0.4533	0.3181	.5727	Form 3 vs. 2: 1.291	0.528	3.154
High prior COS knowledge	1	1.1514	0.4197	7.5259	**.0061****	High prior COS knowledge vs. medium or low prior COS knowledge: 3.163	**1.439**	**7.575**

Note. – = no value provided, not a meaningful value.
** Statistically significant where p < .05.*
*** Statistically significant where p < .01.*

For the 3–5 forms, the full model, including the forms and information about prior COS knowledge, also fit the data significantly better than the reduced model (LRT χ^2= 11.127*, *df* = 3, $p < .05$*; Wald χ^2= 10.1764* = $p < .05$). The form used was not a significant predictor of passing the assessment. However, having high prior COS knowledge was a significant predictor relative to people who were categorized as having low or medium prior COS knowledge

(Table 4). The logistic regression models showed no significant differences in pass rates or difficulty across forms once knowledge level was included, and no evidence of form-by-knowledge interactions.

Table 4. Logistic Regression Results for 3–5 Forms of the COS-KC Assessment (*n* = 138)

Source	*df*	B Estimate	SE B	Wald χ2	*p*	Odds Ratio (OR) Estimate	95% CI OR Minimum	95% CI OR Maximum
Intercept	1	-0.0348	0.3444	0.0102	.9194	–	–	–
Form 4	1	0.1330	0.4469	0.0886	.7660	Form 4 vs. 5: 1.142	0.475	2.764
Form 6	1	0.1120	0.4485	0.0624	.8028	Form 6 vs. 5: 1.291	0.464	2.715
High prior COS knowledge	1	1.2584	0.3947	10.1637	**.0014****	High prior COS knowledge vs. medium or low prior COS knowledge: 3.520	**1.661**	**7.875**

Note. – = no value provided, not a meaningful value.
** Statistically significant where p < .05.*
*** Statistically significant where p < .01.*

Field Test Conclusion

The field-test data substantiate that the COS-KC is a valid assessment tool that can identify individuals with sufficient knowledge to meaningfully participate in the COS process and provide accurate ratings. Together, the fit statistics, reliability, difficulty distributions, targeting evidence, and validation analyses support the final item sets for all six COS-KC forms.

Pilot Test for Usability of the COS-KC System

In addition to validating the adequacy of the COS-KC, the development work also included evaluating the useability of the COS-KC system. The online platform serves as a vehicle to administer the assessment, but it also was designed to allow state and local administrators, with appropriate controls on access, to interact with the registration process and the assessment results. The technical platform functionality includes being able to view a variety of different reports on individuals in an administrator's state or local program who had taken the assessment. The goal of the COS-KC system is to allow administrators at both levels to see individual and aggregate results to identify those who needed more support and identify trends in knowledge of the COS process that should be addressed more systemically.

Purpose

The purpose of the pilot test was to explore the usability and feasibility of the COS-KC system, including the directions, technical features, and reports. Specifically, the team wanted to work with states to:

- Identify and address any technical problems and usability concerns with the platform and system procedures,
- Evaluate and improve the content and the procedures for using the system,

- Evaluate and improve the content and the guidance materials for how to use the system,
- Evaluate the utility and useability of the reports produced by the system, and
- Learn whether the passing/not passing results produced by the system matched their expectations based on their knowledge of the individuals taking the COS-KC.

Methods and Sample

States using the COS process were informed of the opportunity to participate in the COS-KC system pilot test in the spring and summer of 2022, and the pilot test began in the fall. Having already validated the assessment, the results for individuals taking the assessment were considered as legitimate passes (or not). The pilot test was conducted to examine how well the system met the needs of state and local administrators.

The following states, the earliest adopters of the COS-KC, participated in the first phase of the pilot study: Alabama (0 to 3 age group), Alaska (0 to 3), Hawaii (0 to 3), and Wisconsin (3 to 5).

States and local programs began using the COS-KC system gradually over several months. The team maintained ongoing communication with each of these states to identify any issues state or local programs encountered as they began working in the system. Issues deemed significant to the usability of the system were communicated to the Johns Hopkins University Center for Technology in Education, the developer of the online platform, for revision.

Late in 2023, the team began recruiting additional states that were interested in trying out the COS-KC and added several other early adopter states to continue to learn about how to improve the system features for state and local users. The early adopter states that participated in the second phase included Delaware (3 to 5), Maryland (0 to 3 and 3 to 5), South Carolina (0 to 3 and 3 to 5). The team communicated regularly with the early adopter states to gather feedback about what worked well, where there were concerns, and what improvements would be helpful.

Feedback and Improvements

Pilot states and early adopter states were pivotal in helping the team develop procedures for onboarding states to the COS-KC, communicate about the COS-KC to local programs and practitioners, improve COS-KC resources and guidance manuals, improve reports produced by the system, find technical bugs or confusing images within the system, and test out the help desk system. They also helped the team think through the variety of ways that states may adopt the tool, including different levels of involvement in the statewide rollout and different choices about whether or not the tool would be required.

As one additional check on validity, the team regularly monitored the information coming from participating states and listened for concerns about the validity of the assessment. All participating states reported that the numbers of people who passed as well as the specific individuals who passed and did not pass seemed correct given their familiarity with the people and regions. States reported that areas of the state where more COS training had occurred performed better on the assessment.

Pilot Test Conclusion

All of the pilot states and early adopter states continue to use the COS-KC. Their testimony at conferences about the feasibility and benefits of the assessment as well as an ongoing belief in

the usefulness of implementing the tool continues to be a strong measure of value of the COS-KC to states.

Conclusion

The COS-KC was designed to reliably identify individuals with sufficient knowledge of the COS to meaningfully participate in the COS process and reach accurate ratings. The 30 items on each of the six forms of the assessment were developed and validated through an extensive process that involved a national field test with 352 early intervention and early childhood special education providers around the country. The fit statistics, reliability, difficulty distributions, targeting evidence, and validation analyses from the field test support the validity of the final item sets used across all six COS-KC forms.

The pilot test was conducted in two phases with nine EI and ECSE programs from seven states. The study showed that the online platform, including its functionality and reports, are useful and useable for states and local programs. As intended, the system provides insight into what providers in a state know and do not know about the COS process.

References

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Sage Publications.

Konting, M. M., Kamaruddin, N., & Man, N. A. (2009). Quality assurance in higher education institutions: Exist survey among Universiti Putra Malaysia graduating students. *International Education Studies*, *2*(1), 25–31. https://doi.org/10.5539/ies.v2n1p25

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, *16*(2), 878. https://www.rasch.org/rmt/rmt162f.htm

Nunnally, J. C., & Bernstein, I. R. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Pallant, J. (2001). *SPSS survival manual: A step by step guide to data analysis using SPSS for windows* (Version 10). Buckingham Open University Press.

Wilson, M. (2004). *Constructing measures: An item response modeling approach*. Routledge.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. MESA Press.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. MESA Press.

About Us

The contents of this report were developed under a grant from the U.S. Department of Education, #H373Z240001. The contents and resources do not necessarily represent the policy of the U.S. Department of Education, and you should not assume endorsement by the Federal Government. Project Officers: Meredith Miceli and Alexis Lessans.

The DaSy Center is a national technical assistance center funded by the U.S. Department of Education, Office of Special Education Programs. The DaSy Center works with states to support IDEA early intervention and early childhood special education state programs in the development or enhancement of coordinated early childhood longitudinal data systems.

To learn more about the DaSy Center, visit the DaSy Center website at https://www.dasycenter.org/.