Program Guide to Monitoring the Child Outcome Summary Process

The U.S. Office of Special Education Programs (OSEP) requires all state early intervention and preschool special education agencies to report data on three child outcomes: (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.

States use several different approaches to measure child outcomes. Ohio uses the Child Outcomes Summary (COS) process, a team process for summarizing information about a child's progress using multiple sources of information, across settings and situations. Teams apply a seven-point rubric to determine how a child's functioning compares with same-age peers. The accuracy of the ratings documented by teams determines what programs, states, and the Office of Special Education Programming can learn about the effectiveness of early intervention for the children enrolled. Four studies including 36 early intervention programs concluded that when implemented as intended, the COS process produces ratings that are valid for accountability and program improvement purposes (Barton, Taylor, Spiker, & Hebbeler, 2016). The study also found that some programs do not always implement the COS process as intended. Monitoring the data on an annual basis can provide programs with information about how well the COS process is being implemented.

Collecting **high-quality** data (valid and reliable) is a major initiative for Ohio. This guide has been developed to assist the state office in meeting state and federal data monitoring requirements related to Child Outcome Summary data. To that end, this guide contains several sections, including:

1. Why collect Child Outcome Summary Data?
2. Defining-High Quality Child Outcome Summary Data
3. Instructions for Analyzing and Determining the Quality of the Data
4. Frequently Asked Questions
5. Ideas for Ensuring the Quality of Child Outcome Summary Data
6. Additional Resources

Why Collect Child Outcome Summary Data?

In this age of accountability, policymakers are asking questions about the outcomes achieved through participation in programs supported by public funds. Judging the effectiveness of any program requires looking at results, not simply at the process of providing services. The Office of Special Education Programs (OSEP) in the U.S. Department of Education now requires states to report outcome data for children and families served through Part C and Part B Preschool of the Individuals with Disabilities Education Act (IDEA) as part of their Annual Performance Report (APR).

States are required to measure and report on the progress children make between the time they enter a program and the time they exit in each of the outcome areas. Data are to be reported for all children who stay in the program at least 6 months. Specifically, for each outcome, states are to report the percentage of children who:

a. Did not improve functioning
b. Improved functioning, but not sufficiently to move nearer to functioning comparable to same-aged peers
c. Improved functioning to a level nearer to same-aged peers, but did not reach it
d. Improved functioning to reach a level comparable to same-aged peers
e. Maintained functioning at a level comparable to same-aged peers

To report data in these five categories, states must have information about children's functioning at two points in time (entry and exit) and have a way to examine the level of improvement or progress in functioning between those time points. The COS process allows states to report on the percentages of infants and toddlers with Individualized Family Service Plans (IFSPs) who demonstrate improved:

- Positive social-emotional skills
- Acquisition and use of knowledge and skills
- Use of appropriate behavior to meet needs

Validity refers to the accuracy of the data. In other words, to what extent are the COS ratings accurately describing the child's capabilities as compared to his/her same –age peers?

Many states have also begun to use data on child and family outcomes in many different ways to improve their programs. For example, Ohio collect COS data annually to support statewide decision-making and program planning and improvement. In order for the data to be a useful tool for program evaluation and/or planning the data must be accurate, which means it must be **valid** measure of child functioning and must be collected **reliably** over time and across data collectors.

Reliability refers to the extent to which an assessment is consistent. In other words, to what extent do all team members use their knowledge of typical child development to arrive at the same child rating for each of the child outcomes?

Defining High-Quality Child Outcome Summary Data

Researchers studying child outcome summary data look for markers of high quality, including stability of the data over time and consistency of the data among data collection methods. Stability of the data over time can be calculated in two ways, (1) consistency of the entry and exit data for individual children (checking for outliers), and (2) stability of the entry and exit data for the group of children over time (entry and exit correlation). Consistency of the data among varied collection methods is calculated by looking at the correlation between the COS ratings and another standardized tool for evaluating a child's abilities as compared to his/her same-age-peers, such as the Battelle Developmental Inventory or the Bayley Scales of Infant and Toddler Development. This guide describes how to easily perform the calculations for a sample of data using Microsoft Excel and interpret the results in order to determine the quality of the data your state is collecting.

Analyzing and Determining the Quality of the Data

Conducting an appropriate analysis of the data can tell you whether or not your data are high-quality (a reliable descriptions of the child's functional abilities across settings and situations as compared to same-age peers).

Consistency of the Data

Checking for outliers helps us see if any entry-exit pairs of data that have a pattern significantly different from all the other pairs. Outliers are not necessarily "bad data," but they do warrant a closer look. Within COS data, it is common to see entry-exit pairs that are similar (the same at entry and exit or within one to two rating points of one another). If an entry-exit pair shows more than two rating points of growth or decline the pair might be considered suspicious. Sometimes reasonable explanations exist for the differences. For example, consider a premature baby who had a serious medical condition and was hospitalized during the first few months of life when enrollment occurred and over the next two years recovered from the medical condition and demonstrated typical development at exit. This child could reasonably have extremely low ratings at entry and age-appropriate ratings at exit. Typically, though, we would expect that children who are significantly developmentally delayed at entry will continue to be significantly delayed at exit, even when significant developmental progress is made (since at entry and exit they are being compared to same-age peers who are also making significant progress).

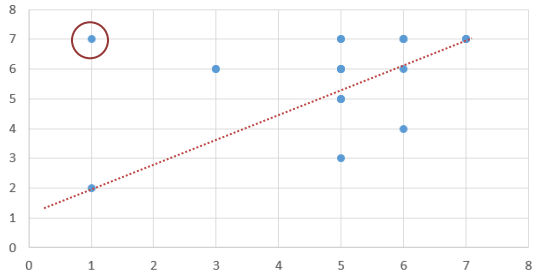
However, if no reasonable explanation exists for the large discrepancy between entry and exit score, it may be that the team needs additional training with using the rubric to rate the child's abilities compared to same-age peers.

One way to check for outliers is to graph the data on a scatter plot. In order to graph the data on a scatterplot, follow these instructions:

- Download a report into Excel that includes the entry and exit scores for each of the three child outcomes for the population of children you want to consider (e.g., state-wide data, count data, regional data, etc.)

- First exclude any lines that do not have both entry and exit data (highlight the line and hide it). Missing data will automatically be converted to a “0” and will appear like an outlier on your scatterplot.
- Highlight an entry column for one of the child outcomes (e.g., social emotional entry) and while holding the control key highlight the corresponding exit column for the same child outcome.
- Click on insert in the menu across the top of your spreadsheet.
- In charts click on the scatter plot.
- The scatter plot will pop up.

This scatter plot shows that the data are fairly closely clumped around an imaginary line that slopes up. This indicates that the scores are positively correlated, meaning that if one score is high the other tends to also be high and if one score is low, the other tends to be low. Although children may have incremental growth in their COS ratings, we would not expect a child’s entrance, annual, and exit data to fluctuate significantly. So, seeing that entry and exit scores are close is a sign of high-quality data.



You may notice an extreme outlier (circled in red above). This chart shows us that one child seems to have been rated a “1” at entry and a “7” at exit. Since this is not very likely, this child’s situation should be looked at more closely to determine whether the rating pair is reasonable or if the team that conducted the COS process needs additional guidance. If the data show multiple outliers, it is possible to determine whether there are patterns or commonalities among the children that are represented by the outlying pair (e.g., the same core team members, or the same service coordinator, multiple or severe disabilities, etc.).

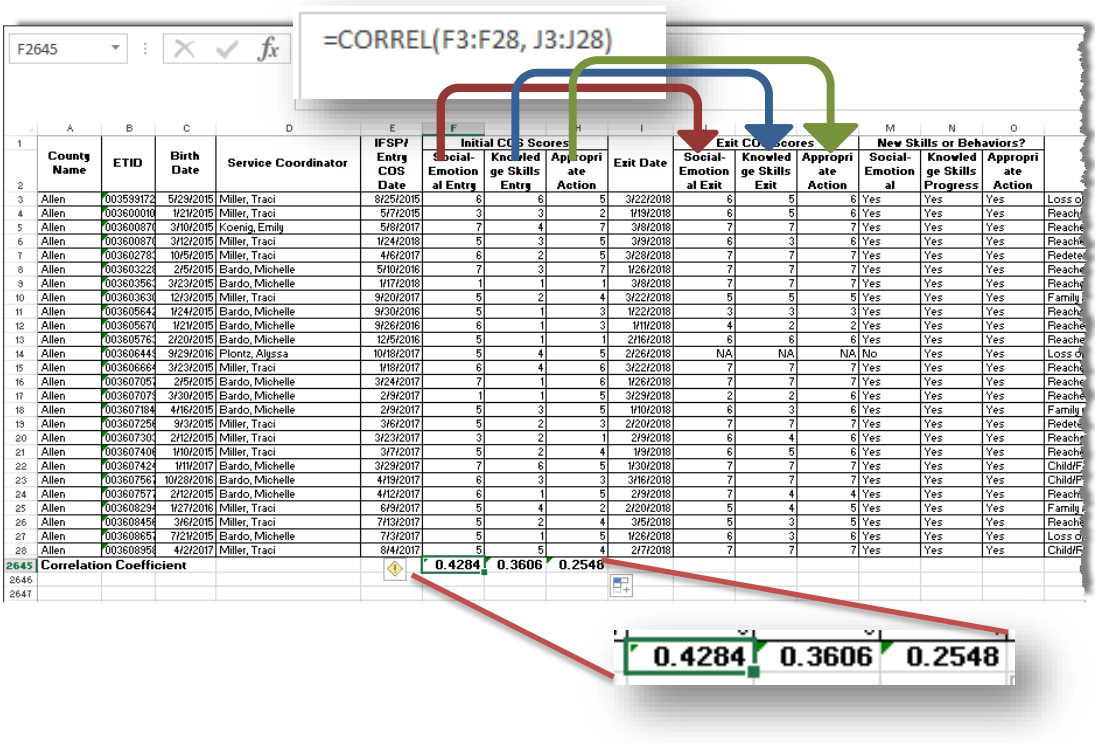
This scatter plot tells us the data are **correlated** and **linear**. Correlated is another way of saying that the data are stable. In other words, children who have high scores at entry also tend to have high scores at exit and children who have low scores at entry tend to also have low scores at exit. Linear means that the data points are clustered in such a way that straight line can be drawn through the data. When the relationship is linear we can perform additional calculations to understand how strongly correlated the pairs of data are (described below).

Stability of Data over Time

When data are linear (on a scatter plot) additional correlation calculations can be performed (e.g. Pearson’s r) to determine how closely correlated the entry-exit pairs are.

In order to test how closely correlated the entry-exit pairs are, you can conduct a correlation test using the report downloaded into Microsoft Excel.

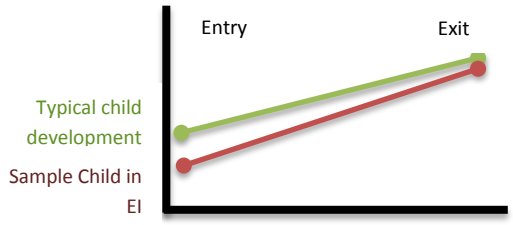
1. Underneath you last line of data in a column labeled with an initial COS outcome, type “Correlation Coefficient”
2. Next we are going to compare the initial COS scores for that outcome with the exit COS scores for the same child outcome. In order to do this type a formula into cell under the last set of data in column F (social emotional at entry) that follows the following format: Correl(F3:F28, J3:J28). In place of “F3” type in the column and row label for your first line of entrance outcome 1 data. In place of “F28” type in the column and row label for your last line of entrance outcome 1 data. In place of “J3” type in the column and row label for your first line of exit outcome 1 data. In place of “J28” type in the column and row label of your last line of exit outcome 1 data. For example, if your report shows data from line 3 to line 250, you will use the formula: Correl(F3:F250, J3:J250).
3. Complete step three for the Knowledge and Skills entry and exit data.
4. Complete step three for the Appropriate Action entry and exit data.



Interpreting the Results

Correlation shows us a pattern of relationship between two sets of numbers. Low correlation numbers (less than .2) indicate that there is very little pattern of consistency between the two sets of numbers. A moderate to high correlation number (above .2) indicates a relationship between the two sets of numbers. In this case:

- A low positive correlation number (less than .2) tells us that there little discernable pattern of relationship between the entry scores and exit scores.
- A negative correlation number (any number with a – sign) tells us that as one number rises, the other number falls. This would mean that COS scores are going down between entry and exit.
- A moderate to high positive correlation number (above .2) tells us that there is a positive (or stable) relationship between the entry and exit scores. Children with disabilities enrolled in early intervention are likely to make progress in development, but even if the child is developing more rapidly than what one would expect based on maturation, when compared to the development of their same age peers, the ratings are likely to be incremental at best. So, (with a few exceptions) the closer the ratings are between entry and exit, the more likely the numbers are to be reliable.



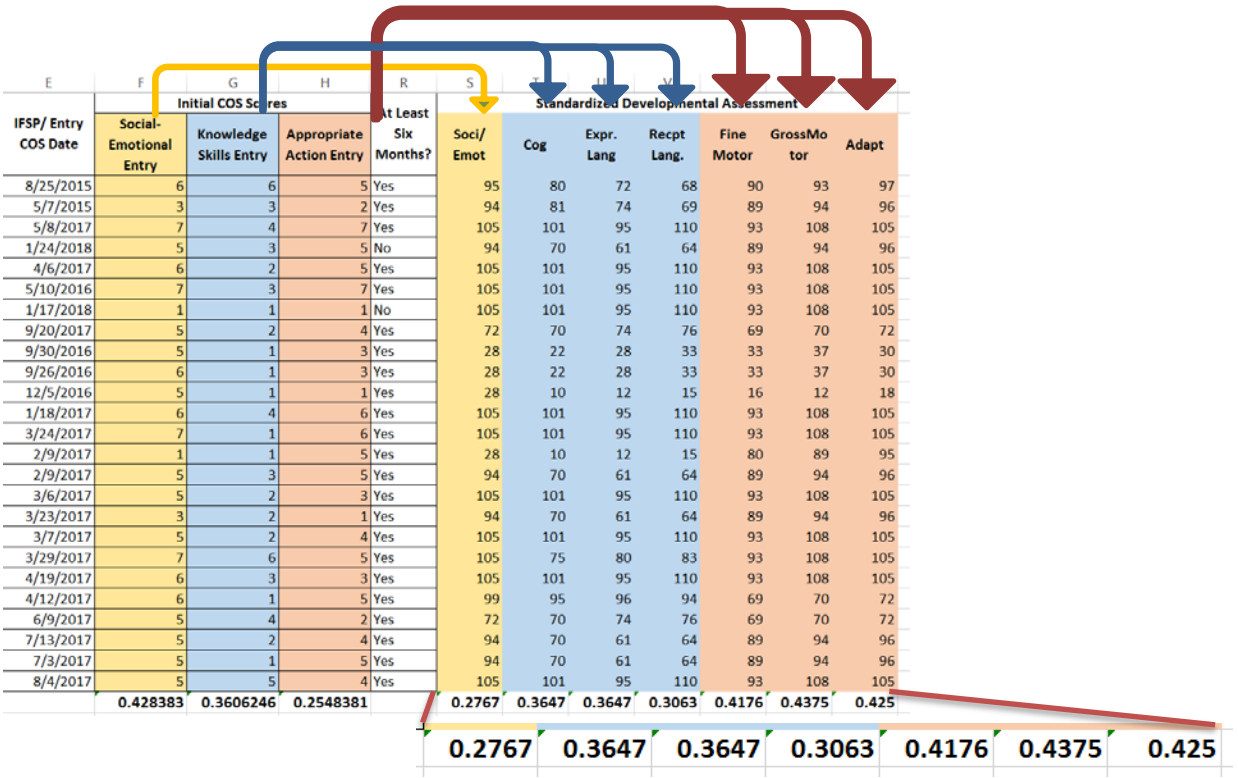
Consistency Over Data Sources

Another indicator of high-quality data is that the data gathered across multiple sources show consistent findings. For example, children who are rated high on the COS also scored high on the standardized assessment tool. Programs can compare eligibility evaluation data or annual reevaluation data to the COS ratings to determine if there is a correlation. COS scores that are strongly correlated with evaluation data may have stronger validity. When state or local program data demonstrate trends of strong correlation among aggregated data, the data are likely to be high-quality. In order to analyze the relationship between multiple data sources, programs can follow the procedures used to correlate entry and exit data above. Data analysts with experience may be able to help you engage in more sophisticated analyses. If you do not have the means to partner with a statistician, you can follow the steps below to employ an effective statistical calculation using Exel functions:

1. Underneath you last line of data for a COS entry outcome (e.g. social emotional entry), type “Correlation Coefficient”
2. Next we are going to compare the initial COS scores with the entry developmental evaluation scores that pertain to each of the child outcomes. In order to do this type a formula into cell under the last set of data in the column that houses the social emotional developmental evaluation data (e.g., column S for the report used in this guide) that follows the following format: Correl(F3:F28, S3:S28). In place of “F3” type in the column and row label for your first line of entrance outcome 1 data. In place of “F28” type in the

column and row label for your last line of entrance outcome 1 data. In place of “S3” type in the column and row label for your first line of standardized evaluation data for the social/emotional subtest. In place of “S28” type in the column and row label of your last line of scores of the standardized evaluation data for the social/emotional subtest. For example, if your report shows data from line 3 to line 250, you will use the formula: Correl(F3:F250, S3:S250).

3. Complete step two for the Knowledge and Skills entry and the cognitive subtest from the developmental evaluation. Complete step three for the Knowledge and Skills entry and the expressive language subtest from the developmental evaluation. Complete step three for the Knowledge and Skills entry and the receptive language subtest from the developmental evaluation.
4. Complete step two for the Appropriate Action entry and the fine motor subtest from the developmental evaluation. Complete step two for the Appropriate Action entry and the gross motor subtest from the developmental evaluation. Complete step two for the Appropriate Action entry and the adaptive behavior scale from the developmental evaluation.



E	F	G	H	R	S						
	Initial COS Scores			At Least Six Months?	Standardized Developmental Assessment						
IFSP/ Entry COS Date	Social-Emotional Entry	Knowledge Skills Entry	Appropriate Action Entry		Soci/ Emot	Cog	Expr. Lang	Recpt Lang.	Fine Motor	GrossMotor	Adapt
8/25/2015	6	6	5	Yes	95	80	72	68	90	93	97
5/7/2015	3	3	2	Yes	94	81	74	69	89	94	96
5/8/2017	7	4	7	Yes	105	101	95	110	93	108	105
1/24/2018	5	3	5	No	94	70	61	64	89	94	96
4/6/2017	6	2	5	Yes	105	101	95	110	93	108	105
5/10/2016	7	3	7	Yes	105	101	95	110	93	108	105
1/17/2018	1	1	1	No	105	101	95	110	93	108	105
9/20/2017	5	2	4	Yes	72	70	74	76	69	70	72
9/30/2016	5	1	3	Yes	28	22	28	33	33	37	30
9/26/2016	6	1	3	Yes	28	22	28	33	33	37	30
12/5/2016	5	1	1	Yes	28	10	12	15	16	12	18
1/18/2017	6	4	6	Yes	105	101	95	110	93	108	105
3/24/2017	7	1	6	Yes	105	101	95	110	93	108	105
2/9/2017	1	1	5	Yes	28	10	12	15	80	89	95
2/9/2017	5	3	5	Yes	94	70	61	64	89	94	96
3/6/2017	5	2	3	Yes	105	101	95	110	93	108	105
3/23/2017	3	2	1	Yes	94	70	61	64	89	94	96
3/7/2017	5	2	4	Yes	105	101	95	110	93	108	105
3/29/2017	7	6	5	Yes	105	75	80	83	93	108	105
4/19/2017	6	3	3	Yes	105	101	95	110	93	108	105
4/12/2017	6	1	5	Yes	99	95	96	94	69	70	72
6/9/2017	5	4	2	Yes	72	70	74	76	69	70	72
7/13/2017	5	2	4	Yes	94	70	61	64	89	94	96
7/3/2017	5	1	5	Yes	94	70	61	64	89	94	96
8/4/2017	5	5	4	Yes	105	101	95	110	93	108	105
	0.428383	0.3606246	0.2548381		0.2767	0.3647	0.3647	0.3063	0.4176	0.4375	0.425

0.2767 0.3647 0.3647 0.3063 0.4176 0.4375 0.425

Interpreting the Results

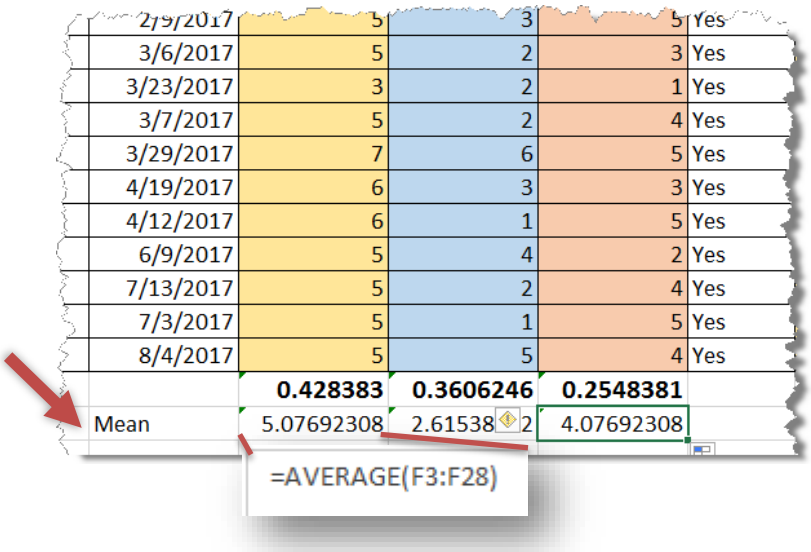
In this example, since the correlation coefficients are all above .2, all of the COS data are strongly correlated with the standardized evaluation data. As the COS data rise, so do the evaluation data that correspond to the outcome area. Likewise, children with lower COS ratings also have lower standardized evaluation data. The standardized evaluation data corroborate the COS ratings. This is a sign that the COS ratings are indeed valid.

Other Ways to Look at the Data

In addition to meeting the OSEP reporting requirement, and testing the quality of the data, child outcome summary data can be used in other ways to inform your program. We can also look at specific characteristics of the data such as the mean (average of the ratings), range (minimum and maximum), mode (rating used the most), standard deviation (how spread out the ratings are) to tell a numerical story about the children served by your program.

Mean

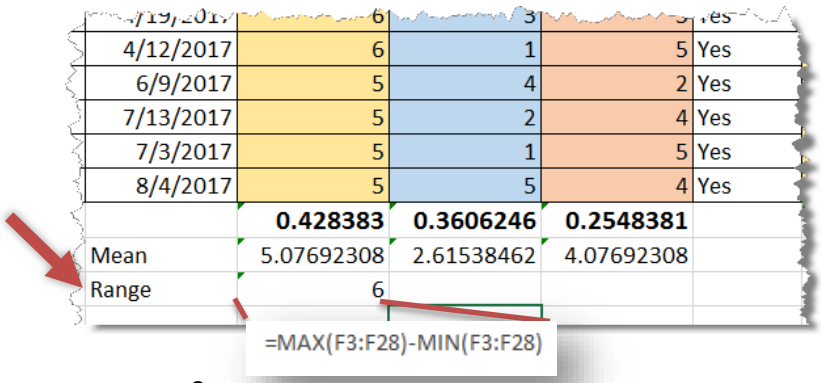
The mean tells you what the average rating are for each global child outcome. The mean can be calculated by clicking on the box below the last line of data in column you would like to average typing Avg. The mean is automatically calculated. Knowing the average tells you about how the population of children you serve compare overall to their same age peers with regards to each global outcome.



2/9/2017	5	3	5	Yes
3/6/2017	5	2	3	Yes
3/23/2017	3	2	1	Yes
3/7/2017	5	2	4	Yes
3/29/2017	7	6	5	Yes
4/19/2017	6	3	3	Yes
4/12/2017	6	1	5	Yes
6/9/2017	5	4	2	Yes
7/13/2017	5	2	4	Yes
7/3/2017	5	1	5	Yes
8/4/2017	5	5	4	Yes
	0.428383	**0.3606246**	**0.2548381**	
Mean	5.07692308	2.61538462	4.07692308	

Range

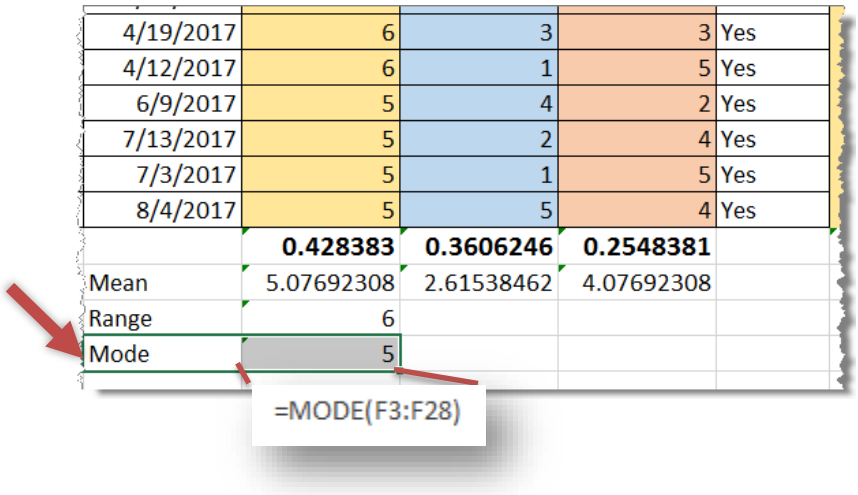
The range tells you the distance between your lowest scores and your highest score for the population of children you serve. Knowing the range can allow you to make statements about the diversity of abilities your program supports. In order to calculate the range, type in a formula that follows the following format: =MAX(F3:F28)-MIN(F3:F28). The result will be the difference between the highest number and the lowest number. In this example, our range tells us that our child outcome ratings for positive social emotional skills spans the entire rating scale from one to sevel. In other words, this program serves children at all developmental levels of social-emotional skills.



4/12/2017	6	1	5	Yes
6/9/2017	5	4	2	Yes
7/13/2017	5	2	4	Yes
7/3/2017	5	1	5	Yes
8/4/2017	5	5	4	Yes
	0.428383	**0.3606246**	**0.2548381**	
Mean	5.07692308	2.61538462	4.07692308	
Range	6			

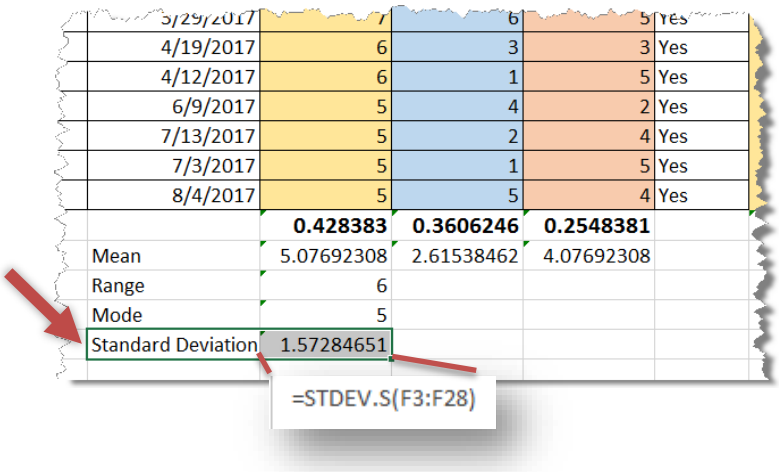
Mode

The mode tells us which number occurs the most times in a set of data. In this case it can tell us which rating is use the most times to describe a specific child outcome among all the children enrolled in your program. In order to calculate the mode, type in a formula that follows the following format MODE(F3:F28). The result will be the rating that is used the most times in the sample of data that falls between lines F3 and F28 in the report. In this case the mode is five (5). In other words, five (5) is the rating used most often by the COS team to describe the positive social emotional skills for the population of children enrolled in the program.



Standard Deviation

Standard deviation tells us how spread out the scores are from the mean (average). A low standard deviation means that the most of the scores are close to the average. A high standard deviation means that the numbers are spread out. For a normal distribution of scores 68% of the scores will fall within one standard deviation of the mean. About 95% of the scores will fall within two standard deviations of the mean. More than 99% of the scores will fall within three standard deviations of the mean. Standard deviation can be calculated by typing a formula that follows the following format: STDEV.S(F3:F28). In this case the standard deviation value tells us that a 68% of the scores fall within 1.5 rating points from the mean and 95% of the scores fall within about three rating points of the mean.



Ideas for Ensuring the Quality of Summary Outcome Data

When you discover that your data are not high-quality (i.e., not stable over time and not consistent over data sources) it is time to employ steps to help those collecting the data become more accurate. An abundance of resources exist to help individuals and teams to work together become proficient at analyzing child outcomes using the COS rubric. We have reviewed many of the available resources and selected and developed a few tools to support teams through each step of the process. The resources and their descriptions can be found at: OhioEarlyIntervention.org

Local programs can use the strategies described below to ensure the data produced by local teams is high-quality.

Monitoring

State and local programs can support state and national efforts by monitoring data locally and ensuring that program staff are able to describe children's functioning with respect to their same age peers. In order to accurately describe and rate the child's level of ability as compared to same-age peers, team members must be able to describe/discussed the following:

- The child's functioning across settings and situations
- Age-expectations for children's development and functioning in the general population
- The content of the three child outcomes
- Guidelines for completing the COS process
- Appropriate age expectations for child functioning within the child's culture

Provide Sufficient Training

Field testing has shown that training is essential to effective and reliable use of the COS process. 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. The Early Childhood Outcomes Center (ECO) recommends 8-12 hours of training for professionals involved in the COS process.

The Early Childhood Outcomes Center (2015) strongly recommends 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.

Frequently Asked questions

How often should we analyze our data?

Since COS scores are updated annually, programs should analyze their data using the procedures described above at least annually. Some programs may choose to conduct the analysis more frequently or conduct different types of analysis on different schedules. For example, a program may decide to disaggregate the data and conduct a separate analyses of the data developed by each team (in addition to the program's aggregated analysis). The team-level analysis may help you understand which teams are collecting high-quality data and which teams need more support.

How important is it that I include all of the data in each analysis?

Depending on the size of your program, analyzing all the data can be prohibitively time consuming. You do not need to analyze all of the data using the procedures above in order to gain a picture of the quality of your data. If you choose not to analyze all of the data available, how you choose which data to include matters and can affect the results of your analysis. The data you analyze must be randomly selected. For example, you should not analyze all the girls, all the children between the ages of 18 and 30 months, all the children with last names A through M, all the children served by a specific team, ten children selected by each team, or use any other pattern. In order for the sample to be random, the selection cannot be based on ANY criteria. One way to ensure a random sample is to assign each child a number and use a random sample generator available for free on the internet to generate a random list of numbers. The random list of numbers will determine which children are selected for analysis.

If I don't analyze all of the data we have collected, how do I decide how many cases to include in the analysis?

When it comes to correlating data, sample-size matters. The larger the sample, the more likely the results are to mirror the entire population. In the case of COS data a random sample of at least 100 children are likely to produce accurate results.

Additional Resources

[Insert the electronic binder resources here]

References

Barton, L., Taylor, C., Spiker, D., & Hebbeler, K. (2016). Validity of the data from the Child Outcomes Summary Process: Findings from the ENHANCE Project. Menlo Park, CA: Center for IDEA Early Childhood Data Systems and Early Childhood Technical Assistance Center.

Early Childhood Outcomes Center (2015). Overview of the child outcomes summary process. Retrieved from http://ectacenter.org/~pdfs/eco/COSF_overview.pdf.

Early Childhood Technical Assistance Center & Center for IDEA Early Childhood Data Systems. (2017). State child outcomes measurement system framework. Retrieved from http://ectacenter.org/eco/pages/childoutcomes.asp#frameworks

Nelson, R., Kelley, G., Hebbeler, K., Vinh, M., Gillaspy, K., Barton, L., & Reid, J. K. (2017). Local child outcomes measurement system (L-COMS). Chapel Hill, NC: Early Childhood Technical Assistance Center. Retrieved from http://ectacenter.org/eco/assets/pdfs/L-COMS_Framework.pdf