

Family Outcome Data: What They Mean and How to Use Them

Batya Elbaum, Melissa Raspa,
Robin Nelson, and Donna Noyes

OSEP Overlapping Data Managers Meeting
Arlington, VA
June 9, 2008

Session topics

- What are “family outcomes?”
- Ensuring the quality of data collected to address Indicator C4
- Interpreting local data
- Developing improvement activities

What are “family outcomes?”

- “We defined ‘family outcome’ as a benefit experienced by families as a result of services received”
 - Bailey, D., Hebbeler, K., Olmsted, M., Raspa, M., & Bruder, M. (2008). Measuring Family Outcomes: Considerations for Large-Scale Data Collection in Early Intervention.
 - Available at http://www.fpg.unc.edu/~eco/pdfs/Measuring_Family_Outcomes_revised_27-April.pdf

Part C State Performance Plan Indicator #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.”

Requirements of the SPP

- The data for Indicator C4 must come from families
- States may use any method of data collection; most, if not, all states are using a survey approach
- States are free to select or create their data collection instrument
- Whether the distribution of surveys is to the whole population of families participating in Part C or to a sample of such families, states must report on the representativeness of their obtained sample

Comparison of the ECO and NCSEAM instruments in relation to **content**

- The ECO Family Outcomes Survey and the NCSEAM Impact on Family Scale have similar content
- The instruments were developed separately by ECO and NCSEAM, but both centers used significant input from families and other stakeholders to develop instrument content
- The content of both instruments goes beyond the content of the three indicators specified in Part C Indicator 4

Comparison of the ECO and NCSEAM instruments in relation to **items**

- The ECO Family Outcomes Survey addresses five outcome areas, each of which is represented by 3 items.
- There are three additional questions that map directly onto the SPP indicators.
- The NCSEAM Impact on Family Scale draws on a potentially unlimited set of validated items that relate to the overall construct of “family outcomes.” NCSEAM has recommended that a minimum of 22 items be used.

Comparison of the ECO and NCSEAM instruments in relation to **focus**

- The ECO Family Outcomes Survey focuses on families' perceptions of **their current status** in relation to achieving positive family outcomes
- The NCSEAM Impact on Family Scale focuses on families' perceptions of the extent to which **early intervention services helped them** to achieve positive family outcomes

NCSEAM

- National Center for Special Education Accountability Monitoring
- Developed measurement scales for states' use in addressing C4 and B8

What is the “Part C NCSEAM Survey?”

- What most people mean when they refer to the “Part C NCSEAM Survey” is the Impact on Family Scale (IFS).
- The IFS measures the extent to which, in the eyes of the family, early intervention services have helped the family achieve positive family outcomes.

The IFS

- NCSEAM validated a set of items related to family outcomes as part of its National Item Validation Study involving over 1750 families who had participated in early intervention. Items validated by NCSEAM are in the “Part C Item Bank” available on the web at www.accountabilitydata.org.

The IFS

- NCSEAM recommended that states use a minimum of 22 validated items in order to ensure reliabilities of .90 or above.
- For each item, respondents are asked to choose one of 6 responses ranging from “very strongly disagree” to “very strongly agree.”

Measurement framework

- NCSEAM applied the same measurement framework used in the development and validation of assessment tools such as the SAT, GRE, LSAT, MCAT, psycho-educational assessments, developmental assessments, and statewide achievement tests in reading and mathematics

Simple test of math skills

- Item . $\sqrt{578}$
- Item . $1 + 1$
- Item . 4×56
- Item . $(x^2 + 4)(x - 5) = 52$
- Item . $374 \div 19$

Item Order

- Item 1. $1 + 1$
- Item 2. 4×56
- Item 3. $374 \div 19$
- Item 4. $\sqrt{578}$
- Item 5. $(x^2 + 4)(x - 5) = 52$

Item Spacing

• 1. 2. 3. 4. 5.

• 1. 2. 3. 4. 5.

Items families agree with most

Over the past year, early intervention services have helped me and/or my family:

- Do things with and for my child that are good for my child's development.
- Understand my child's special needs.
- Be more optimistic about my child's future.
- Feel that my efforts are helping my child.

Items families agree with *less*

Over the past year, early intervention services have helped me and/or my family:

- Communicate more effectively with the people who work with my child and my family.
- Know about my child's and family's rights concerning special education services.
- Understand the roles and responsibilities of the people who work with my child and family.

Items families agree with *much less*

Over the past year, early intervention services have helped me and/or my family:

- Know where to go for support to meet my child's needs.
- Know about services in the community.
- Participate in typical activities for children and families in my community.

The IFS measures: What do the numbers mean?

- The numbers we use to measure things are arbitrary. For the IFS, NCSEAM picked a scale that is similar to the SAT scale. Interpretable measures go from 200 to 800.
- Individual measures are expected to vary. Some will be very low (200-300), some will be very high (700-800), and most will be in somewhere in between.

Going from the distribution of measures to the percent for 4A, 4B, and 4C

- We often measure a phenomenon on a continuous scale but then use a decision rule that determines which individuals are “proficient,” “healthy,” etc. The cut-score that divides the distribution is often referred to as the “standard.”

Going from the distribution of measures to the percent for 4A, 4B, and 4C

- For purposes of reporting on Indicator C4, we need to determine the measure at which we can say that families at or above that measure “report that early intervention helped them to achieve” the three specific outcomes.

How high does the IFS measure have to be for 4A?

- A national stakeholder group convened by NCSEAM reached the consensus that the minimum measure for 4A should be 539.
- We know that at a measure of 539 or above, families are expressing **confident agreement** that early intervention services have helped them know their rights.

How high does the IFS measure have to be for 4B?

- The same national stakeholder group reached the consensus that the minimum measure for 4B should be 556.
- We know that at a measure of 556 or above, families are expressing **confident agreement** that early intervention services have helped them effectively communicate their children's needs.

How high does the IFS measure have to be for 4C?

- The same national stakeholder group reached the consensus that the minimum measure for 4C should be 516.
- We know that at a measure of 516 or above, families are expressing **confident agreement** that early intervention services have helped them help their child develop and learn.

How do we know the instrument yields valid data?

- The content was generated by families and other knowledgeable stakeholders and affirmed by family organizations and researchers
- Statistical analyses indicate that the items all work as a single scale (i.e., they do not separate out into subscales)

How do we know the instrument yields reliable data?

- To address the question of reliability, we need to distinguish two types of reliability that apply to use of the scale in the context of the SPP
 - Reliability of individual measurements
 - Reliability of the estimate of a population percentage

Reliability of individual measurements

- To what extent can we be confident that the same instrument or procedure, applied to the same person, would yield the same result if the measurement were repeated on two occasions very close in time, or if the measurement were taken by different individuals?

Quantifying the reliability of individual measurements

- How reliable is an individual measurement?
- All measurements contain a certain amount of error (unreliability) because no measurement instrument is perfect (measurement error/tool) and no user uses a measurement instrument perfectly (measurement error/user).

Quantifying the reliability of individual measurements

- Statistics used to express measurement reliability range from 0.00 (no consistency) to 1.00 (perfect consistency).
- If the reliability is high, for example, .90 or above, the measurement has little error and is highly reliable.

What high reliability buys you

	Separation Reliability	Number of statistically distinct ranges
•	.50	1
•	.70	2
•	.80	3
•	.90	4
•	.94	5
•	.96	6
•	.98	7
•	.99	8+

Reliability is not a property of an instrument

- Reliability is a function of the interaction of the instrument and the user
- Therefore, reliability estimates reported by the developers of an instrument may or may not be replicated by other users
- Best practice requires reporting the reliability of the instrument based on each group of persons measured

Reliability of the estimate of a population percentage

- To what extent can we be confident that the same instrument or procedure, used with a different sample drawn from the same population, will yield the same result?

Quantifying the reliability of an estimate of a population percentage (“Percent of parents who report that ...”)

- How reliable is an estimate of a population percentage, based on a sample percentage?
- All estimates based on a sample contain a certain amount of error because they do not have information from the whole population (sampling error).

Quantifying the reliability of an estimate of a population percentage (“Percent of parents who report that ...”)

- Statistics used to express the reliability of an estimate of a population percentage focus on the amount of unreliability, or error, in the estimate. Two examples are confidence intervals and margins of error.
- If the margin of error is small, for example, $\pm 1\%$, the estimate has little error and is highly reliable.
- As sample size goes up, sampling error goes down.

Reporting at the local program level

- What does OSEP require?
- What do states want to share with their local programs and the public?
- What makes sense, given the data?

Family Outcome Data – What Do They Mean and How To Use Them

New York State GSEG – A Model Approach to Program Evaluation and Quality Improvement

OSEP Overlapping Data Meeting

June 11, 2008

Identifying Child and Family Outcomes Using Concept Mapping

Concept Mapping Process

Planning: Project staff and Core Group developed 2 *focus prompts* - "*As a result of early intervention services, children (families) will...* and identified *participants (245)*

Idea Generation: Participants **brainstormed ideas**, knowledge and opinions on a **project website** – 2091 ideas! *Planners synthesized these ideas to manageably sized set of 119 ideas*

Organizing: *all participants rated ideas* on **importance and impact**, and core advisory group **sorted ideas into conceptually similar groups**

Analysis: CSI ran analysis, created *maps*, then *pattern matches and "go zones,"* based on the map

Interpretation: planners, core advisory group, and staff met to interpret the results and identify potential measurement strategies

What We Learned through Concept Mapping

- High level of interest and involvement by stakeholders across the state
- High level of agreement by stakeholders across the state on the potential impact of EI
- Child outcomes and family outcomes are conceptually different
- A distinct set of child and family outcomes for NYS could be identified

Measuring Child and Family Outcomes – The Field Study

Field Study Goals

- Identify ways to measure child and family outcomes identified by NYS stakeholders
- Design an approach to make full use of NYS concept mapping data set
- Identify potential measurement strategies that could be used for program evaluation AND quality improvement

Field Study Approach

- The National Center for Special Education Accountability Monitoring (NCSEAM) “Impact of EI on the Family” national scale served as the basis for developing a NYS “Impact of EI on the Family” scale

NYS Family Survey Development

- Child and Family Outcomes items identified using concept mapping formed the basis of two scales
- Outcome items reworded to ensure clear, simple, unambiguous and readable items
- Crosswalk between NYS concept mapping items to NCSEAM national Impact of EI on the Family item bank
- 158 items in the final survey

Field Study Participants

- **515** families of children exiting the EIP in Nassau and Suffolk Counties in the Spring, 2007
- **230 completed the survey** (a response rate of **45%**)
- **12** parents also participated in **90 minute telephone focus group interviews**
- **58** parents completed **feedback forms** about the survey

Parent Feedback on the NYS Family Survey

Parent Feedback on Uses

Parents in the study recommended:

- Using the survey as checklist for parents as they enter the EIP to help them:
 - Learn the EI “lingo”
 - Participate more fully in goal setting
 - Be less skeptical of assessment
- Administering the survey routinely through the EIP as a “goals check”
- Administering the survey 3-6 months post EIP as a follow-up to see how children and families are doing/how EI helped

Family Survey Results

Rasch Measurement Model Analyses – Summary of Findings

- The NYS scales are highly reliable and unidimensional
- The distribution of person measures is normal for both scales
- Principal components analyses align with the concept mapping process in supporting the separation of child and family outcomes

Family Survey Results

- Impact on Family Scale items in order of calibration (from word document)
- Impact on Child Scale Items in order of calibration (from word document)

Interpreting and Using the Data

Using Data with Core Advisory Group

- Results presented/discussed with Core Advisory Group in 2-day meeting
- General comments/questions included:
 - Impressive response rate
 - “Parent satisfaction” vs. Impact of EI on Child and Family Outcomes
 - How representative are LI data?

Core Advisory Group Discussion

- Observations
- NYS-specific outcomes
- Standard setting
- Potential Uses for Family Impact Scale
- Child Impact Scale

Core Advisory Group Discussion

Recommendations

- Collect more data
- Integrate the scales into multiple levels of the EI Process
- Find ways to broadly disseminate the results
- Present and discuss the results with key stakeholders in Nassau and Suffolk Counties
- Keep the core advisory group informed/involved

Using Data with Local Programs

- Joint meeting of Nassau and Suffolk County LEICCs on January 24, 2008
- Presentation and discussion about the data
- Participants asked thought-provoking questions

Using data with Local Programs

- Participants excited to have outcome data
- High level of response from families shows families want to have a voice
- Interesting and insightful observations were shared and discussed

Local Programs Discussion

- Productive discussion about what the data might mean/say about their programs
- Brainstorming on how to use the data

Local Programs Discussion

Local Program Recommendations

- Communicate results to families in the study
- Find local resources to use the scales
- Use as tool to monitor individual progress
- Incorporate the family survey into the State EI monitoring process
- Use results to inform the public and political leaders about the positive impact of the EIP

Project Evaluation – Recommendations and Next Steps

- Consider adding outcomes to NYS reporting, particularly related to knowledge of or use of community services and community integration
- Further development of the child outcomes scale
- Create process for translating outcome results into program planning and improvement

Project Evaluation – Recommendations and Next Steps

- Create feedback loops to parents, describing the difference their input makes
- Consider incorporating findings into training programs for EI providers
- Develop, train on and test a tool for use in the IFSP process based on the items
- Disseminate results broadly