

Validating the Child Outcomes Summary Form (COSF) for Use in Accountability Systems for Programs Serving Young Children with Disabilities

Charles Greenwood¹, Dale Walker¹, Margy Hornback², Marybeth Wells³, Kathleen Hebbeler⁴, Donna Spiker⁴, Lauren Barton⁴, Sangeeta Mallik⁴,
¹Juniper Gardens Children's Project, University of Kansas, ²Kansas State Department of Education, ³Idaho State Department of Education, ⁴SRI International

Presented at CRIEI, San Diego, February 2008

Abstract

The Child Outcomes Summary Form (COSF), developed by the Early Childhood Outcomes Center (ECO), is in widespread use in state accountability systems for young children with disabilities. These systems have not been fully validated. Responsive to both discipline and policy concerns (e.g., accommodate multiple sources of measurement, aggregation of data on child functioning across states, etc.), the COSF was developed as a workable solution. We describe propositions that support a scientific argument for the validity of the COSF used in state accountability systems and present data in preliminary support for each claim. Discussion focuses on the validity propositions put forth, the strength of preliminary evidence provided, and future research needed to support validity.

Child Outcomes Summary Form (COSF)

- Developed to support states with federal reporting requirements and for program accountability – child outcomes for young children with disabilities or other special needs
- Developed to allow local teams to synthesize information on child functioning across settings into a 7-point rating
- Timeline of OSEP's requirements did not allow for advance field testing
- Now in use in:
 - 33 state early intervention systems
 - 27 state early childhood special education systems
- Allows aggregation of
 - Multiple sources of information about a child
 - Information from multiple settings
 - Information from multiple assessment tools (either on a single child or from various programs within a state)
 - Data from developmental domains into 3 functional, global outcomes as required for reporting by OSEP

Validity

- Validity and reliability are characteristics of a set of scores derived from an assessment (or an assessment summary process), not characteristics of the assessment itself (APA/AERA/NCME, 1999; Thompson, 2003)
- A case for validity requires (AERA/APA/NCME Standards, 1999):
 - Evidence and theory supporting the interpretation of test scores for their proposed uses
 - A set of propositions with evidence that support the proposed interpretation of scores.

Validity of a State Accountability System

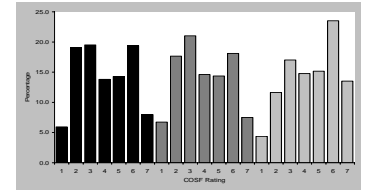
- An accountability system can be said to have validity when evidence is judged to be strong enough to support inferences that:
 - The components of the system are aligned to the purposes, and are working in harmony to help the system accomplish those purpose
 - The system is accomplishing what was intended (and not what was not intended) (Marion et al., 2002, pg. 105)
- Because of the high-stakes nature of accountability systems, there is an obligation to demonstrate the validity of those systems for the purposes for which those data are used (CCSO, 2004)

COSF Validity Argument: Claims, Rationale, and Evidence to Date

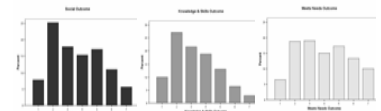
Claim: If the COSF is producing valid information, we would expect that:	Theoretical/Empirical Rationale	Evidence to Date
1. The distribution of COSF ratings at program entry to be sensitive to individual differences (e.g., fewer scores at the high and low ends of the distribution, more scores in the middle).	Available evidence suggests EI and ECSE serve more mildly than severely impaired children (few ratings of 1). Few children receiving services would be expected to be considered as functioning typically (few ratings of 6 or 7).	<i>Supported.</i> See Figure 1. Pattern observed across multiple states.
2. The majority of children show an entry rating of below 6 in at least one of the 3 outcome areas at entry.	Few children receiving Part C or B are expected to be functioning at age expected levels at entry into EI or ECSE.	<i>Supported.</i> State 1: 76%, 70%, and 69% of children score below 6, on outcomes 1, 2, and 3 respectively (Part C and B combined)
3. COSF entry ratings should produce similar distributions across states.	The overall rationale for claim #1 applies to each state. Even the smallest state serves more than 600 children which is enough to produce a stable distribution of ratings.	<i>Supported.</i> Data analyzed from three Part C states two Part B states show similar patterns. Median ratings in all states for all outcomes are 4 or 5 (Figure 1).
4. COSF entry ratings should produce similar distributions across programs with large numbers within states given that programs are serving children with relatively similar characteristics.	The overall rationale in claim #1 applies at the program level if the programs are known to be serving similar types of children.	<i>Not supported.</i> Distribution of ratings for programs within a state were highly variable (one state, six programs with N > 40). Programs were early adopters and all users may not have received consistent training.
5. COSF entry ratings should show moderate but not overly strong relationships with standardized assessment measures used.	COSF ratings are based on a synthesis of case information, only partly from anchor assessment measures. Thus, anchor measures should correlate only moderately to COSF ratings.	<i>Supported.</i> Relationship between COSF rating and child's score on the BDI has been examined. Correlation = .61 to .65 across outcomes (one state; N = 122 with Battelle data).
6. Some child characteristics should be irrelevant to COSF ratings, while others may be salient.	If who receives services does not differ by gender, race, etc., outcome ratings on the outcomes should be similar.	<i>Supported.</i> See Figure 2. Additional data for child's race/ethnicity (one state).
7. Distributions of entry ratings should be different for children with different disabilities.	For some children, the nature of the disability will be related to the child's level of functioning and should be reflected in the COSF ratings.	<i>Supported.</i> Ratings for children with developmental delays (Median = 3 for outcomes 1, 2, and 3; N = 624) were lower than for children with speech language impairments (Median = 5, 4, 5 for outcomes 1, 2, and 3 respectively; N = 296) (one state).
8. Children with certain types of disabilities will have different score patterns across the outcomes compared to children with other types of disabilities.	Some disabilities affect some areas of functioning more severely than others (e.g., children with autism lower ratings for outcome 1 compared to the other outcomes; Children with Down Syndrome lower ratings in outcome 2).	<i>Supported.</i> See Figure 2. (State 1)
9. There should be moderate but not overly high relationships among the three outcomes.	For many, but not all, children with disabilities, functioning in the three outcomes proceeds together.	<i>Supported.</i> Correlations ranged from .62 to .80 for states with N > 1,000. Correlations ranged from .5 to .80 across five state programs.
10. COSF ratings should not be related to the type of anchor assessment used (if providers are not choosing assessments based on the level of child functioning).	An assumption underlying the COSF is that functioning can be validly described across multiple sources regardless of the nature of those sources.	<i>Mostly supported.</i> Ratings were found to be similar with most assessments. More investigation needed about who receives which assessments (two states).
11. COSF ratings should be sensitive to improvement over time following a history (at least 6 months) of EI or ECSE services.	Programs are pursuing recommended and evidence-based practices in the services provided.	<i>Supported.</i> Exit distributions are more positively skewed than the Entry distributions. Mean/Median scores in all outcomes were higher at Exit (two states – see Figure 3).
12. Entry and exit COSF ratings should be moderately correlated.	Some children display some stability in their development over time, e.g., with lower functioning children gaining slower. Alternately, some children should not follow this pattern and make substantial change.	<i>Supported.</i> Entry to exit correlations for State 1: $r = .48$ to $.61$ over outcomes (N = 106); State 2: $r = .71$ to $.74$ over the 3 outcomes (N = 100). These samples were for very early exits and are not fully representative of the entire program.

Note: Based on data for 3 state Part C programs (N = 51; N = 108, N = 1,108) and 2 state Part B Preschool Programs (N = 1,280; N = 1,103).

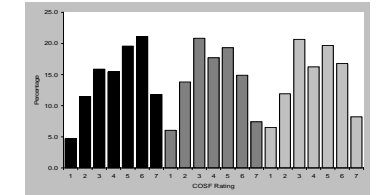
Figure 1



State 1: ECSE program entry ratings for 3 outcomes

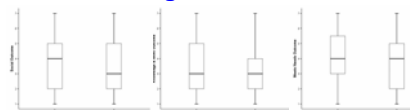


State 2: ECSE program entry ratings for 3 outcomes

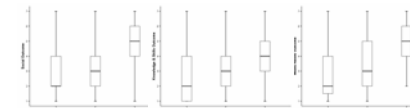


State 1: EI program entry ratings for 3 outcomes

Figure 2

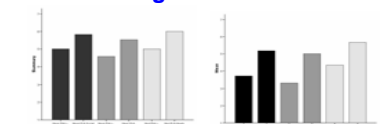


State 2: State ECSE outcomes by gender were similar except that boys were lower on the social outcome



State 2: State ECSE program outcome differences by disability status: (other, developmental delay, speech/language)

Figure 3



State 1 & 2: State ECSE program improvement in mean COSF ratings by outcome

Discussion

- These data appear to provide initial support for use of the COSF in state accountability.
- However, research of this nature is needed to support the validity claims of each state