

Comparing Child Outcomes Summary Ratings to Scores from Assessment Tools

Lauren Barton, Cornelia Taylor, Kathleen Hebbeler, & Donna Spiker

Measuring and Improving Child and Family Outcomes Conference
Minneapolis, MN October 2012

SRI International



Background

- The Child Outcomes Summary (COS) process is the basis for accountability data in more than 40 state early intervention (EI) or early childhood special education (ECSE) programs.
- The COS process was adopted by states to meet federal requirements before studies validating the COS had been conducted.

What is ENHANCE?

- The first large-scale project designed to investigate the validity of the COS process
- The main research question is: Are COS data valid for accountability purposes?
- A set of studies examining validity—a characteristic of the data produced and how they are used; not a characteristic of the tool
- Project approach involves:
 - Developing a set of 17 claims expected to be true if data are valid
 - Testing those claims using multiple approaches in many different programs
 - Four studies that are underway in 36 sites
- Data presented are:
 - Preliminary (based on about 2/3 of the final projected sample)
 - From one study—child assessments study
 - Focus on only one of the 17 validity claims under investigation

Instruments

Child Outcomes Summary (COS) Process

- Teams synthesize multiple sources of information about how children function across settings.
- Teams consider children's functional skills integrated across domains.
- Teams assign a consensus rating about the child's current functioning relative to age-expected functioning for each outcome.
- Ratings range from a low of 1 to a high of 7, with specific criteria for each.
 - 6 & 7 reflect functioning within age expectations.
 - 1-5 reflect increasing distance from age-expected levels of functioning.
- Progress is determined based on the difference between entry and exit ratings and attainment of new skills between entry and exit.

Outcomes measured with COS process:

- Children have positive social relationships.
- Children acquire and use knowledge and skills.
- Children take appropriate action to meet their needs.

Child Assessments Study—Initial Sample

Sample	EI	ECSE
Number of children	71	49
Mean age at entry (SD)	15.3 months (9.8)	43.5 months (7.7)
Age at entry	31% < 8 months 31% 8-20 months 38% > 20 months	8% < 36 months 61% 36-47 months 31% > 48 months
Male	64%	59%
Disability	75% developmental delay 25% diagnosed condition	40% speech/language impairment 38% developmental delay 22% another diagnosis, most frequently autism
Number of states	7	6
Number of programs	19	14

NOTE: These are preliminary data for only a portion of the final projected sample.

Child Assessments Study—Sample COS Ratings

Figures 1 & 2. Percentage of children with each COS rating, by outcome

Figure 1. EI (n = 71)

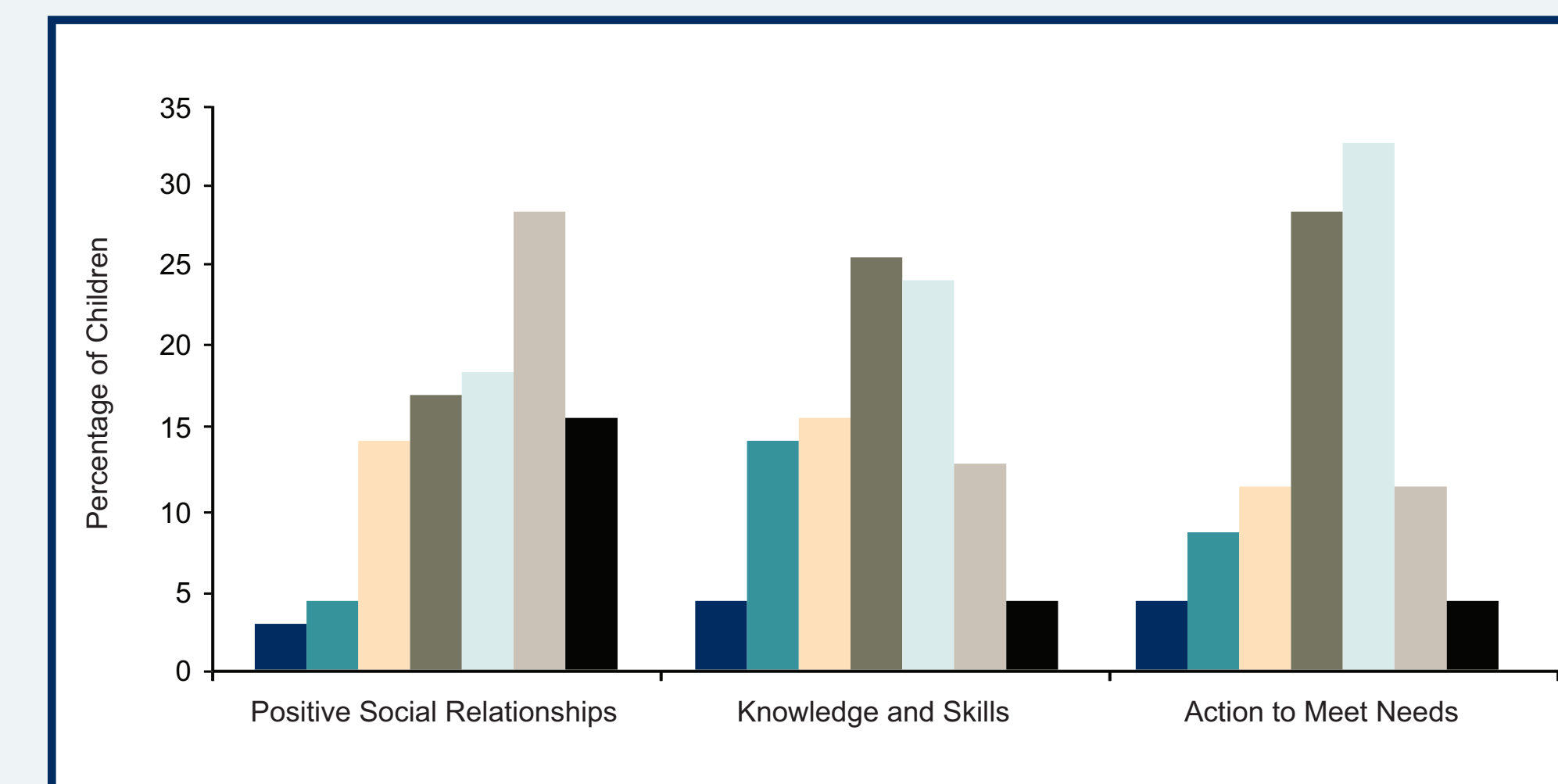
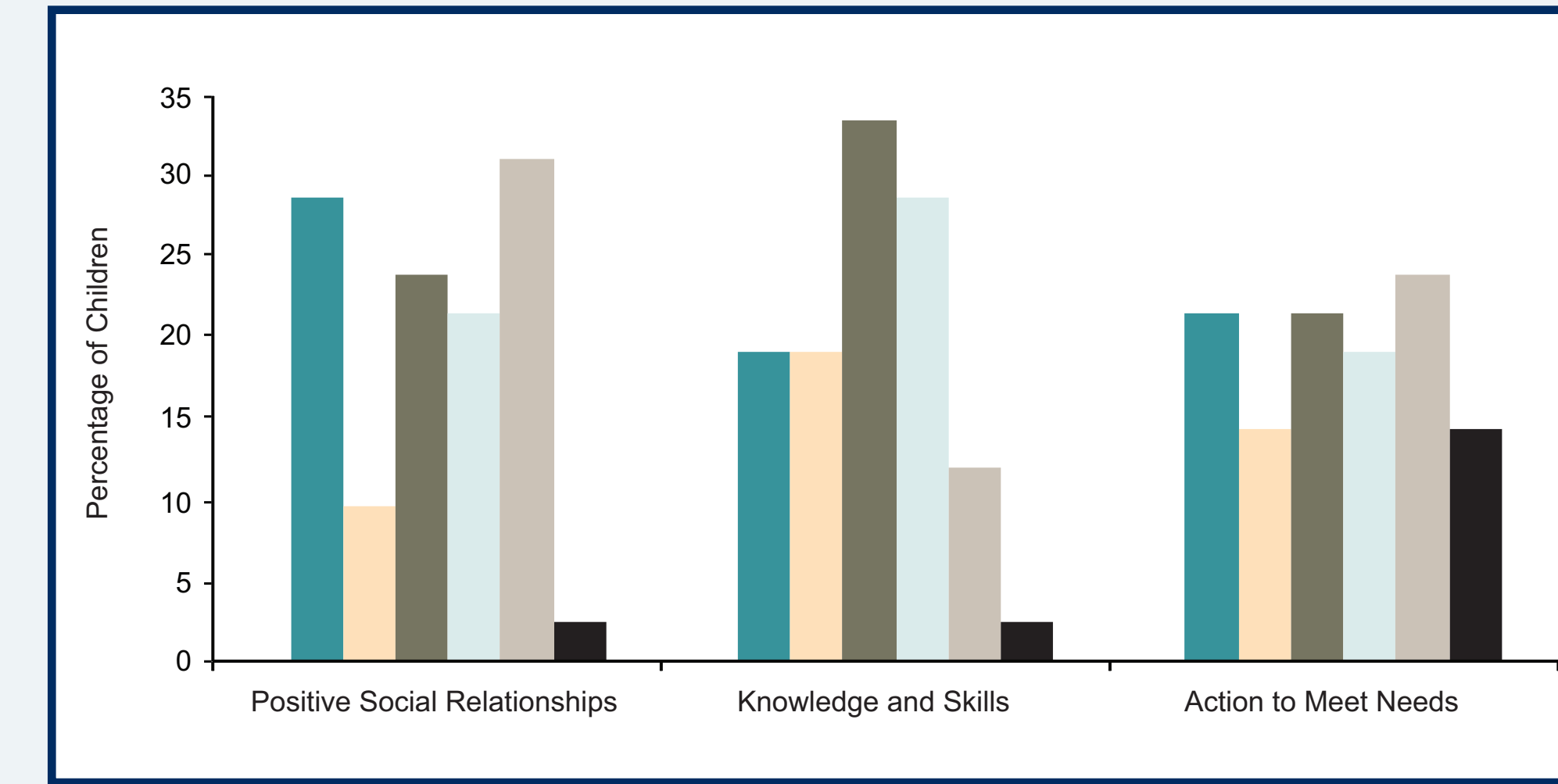


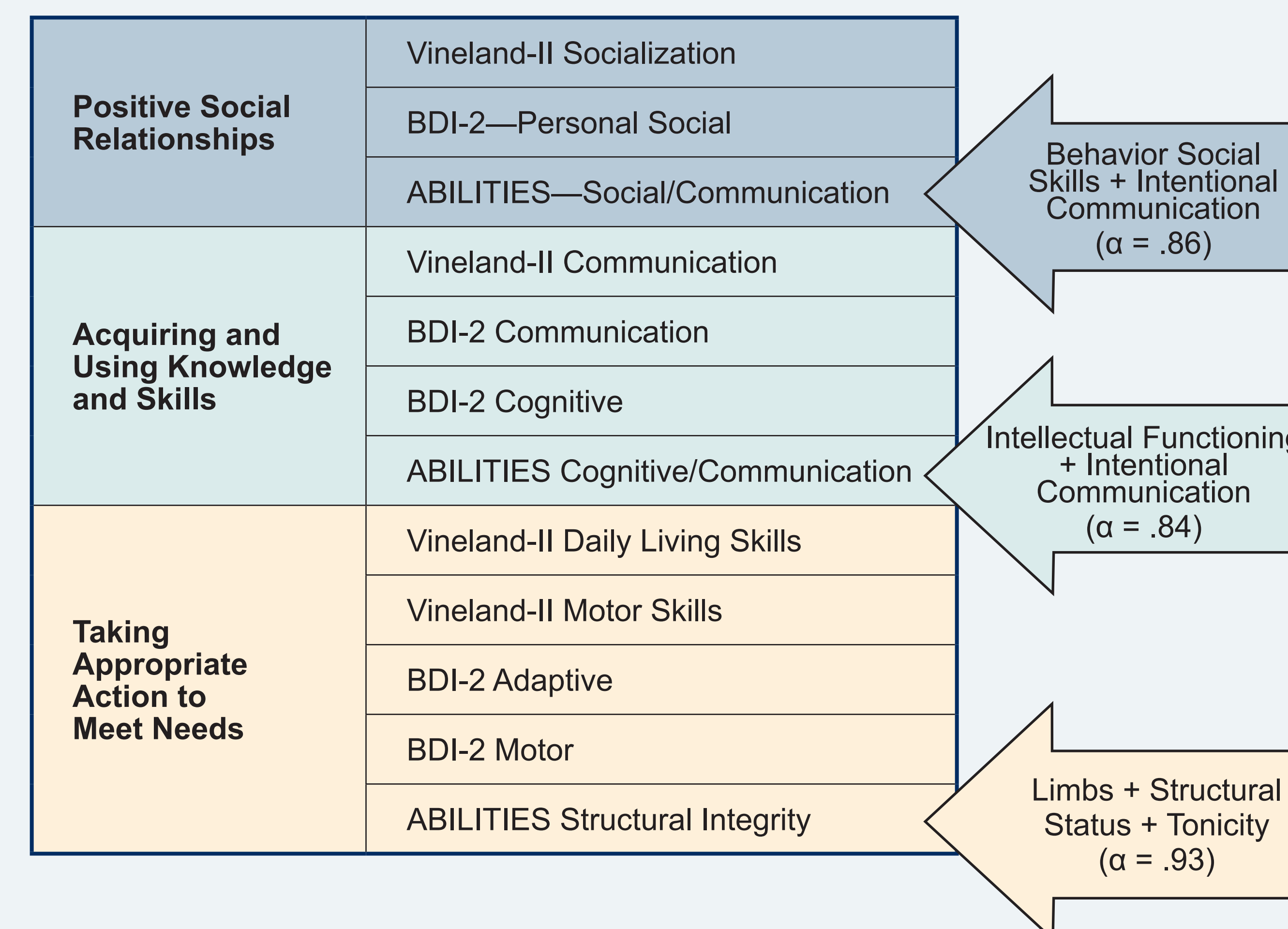
Figure 2. ECSE (n = 49)



Ratings: 1 (dark blue), 2 (teal), 3 (orange), 4 (brown), 5 (light blue), 6 (grey), 7 (black)

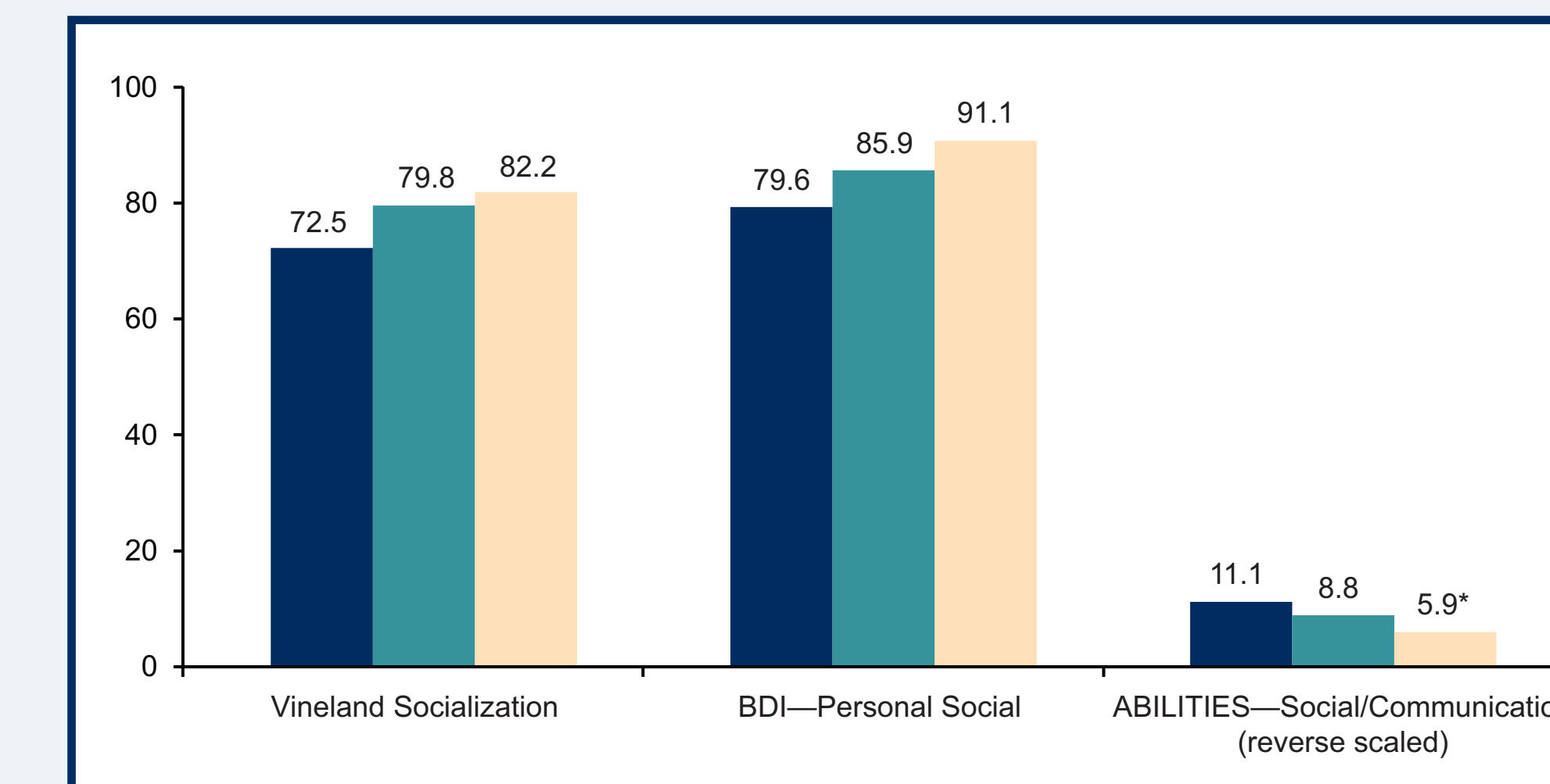
Methods

- Mapped assessment tool domains to COS outcomes.
- Grouped COS ratings into three categories:
 - COS Group 1: Ratings 1, 2, and 3 (foundational skills)
 - COS Group 2: Ratings 4 and 5 (some age expected skills)
 - COS Group 3: Ratings 6 and 7 (at or above age expectations)
- Used one-way ANOVA to compare mean scores on assessment tool domains for each COS group.
- Tested pairwise comparisons between each of the three COS groups, with a Bonferroni correction for multiple comparisons.



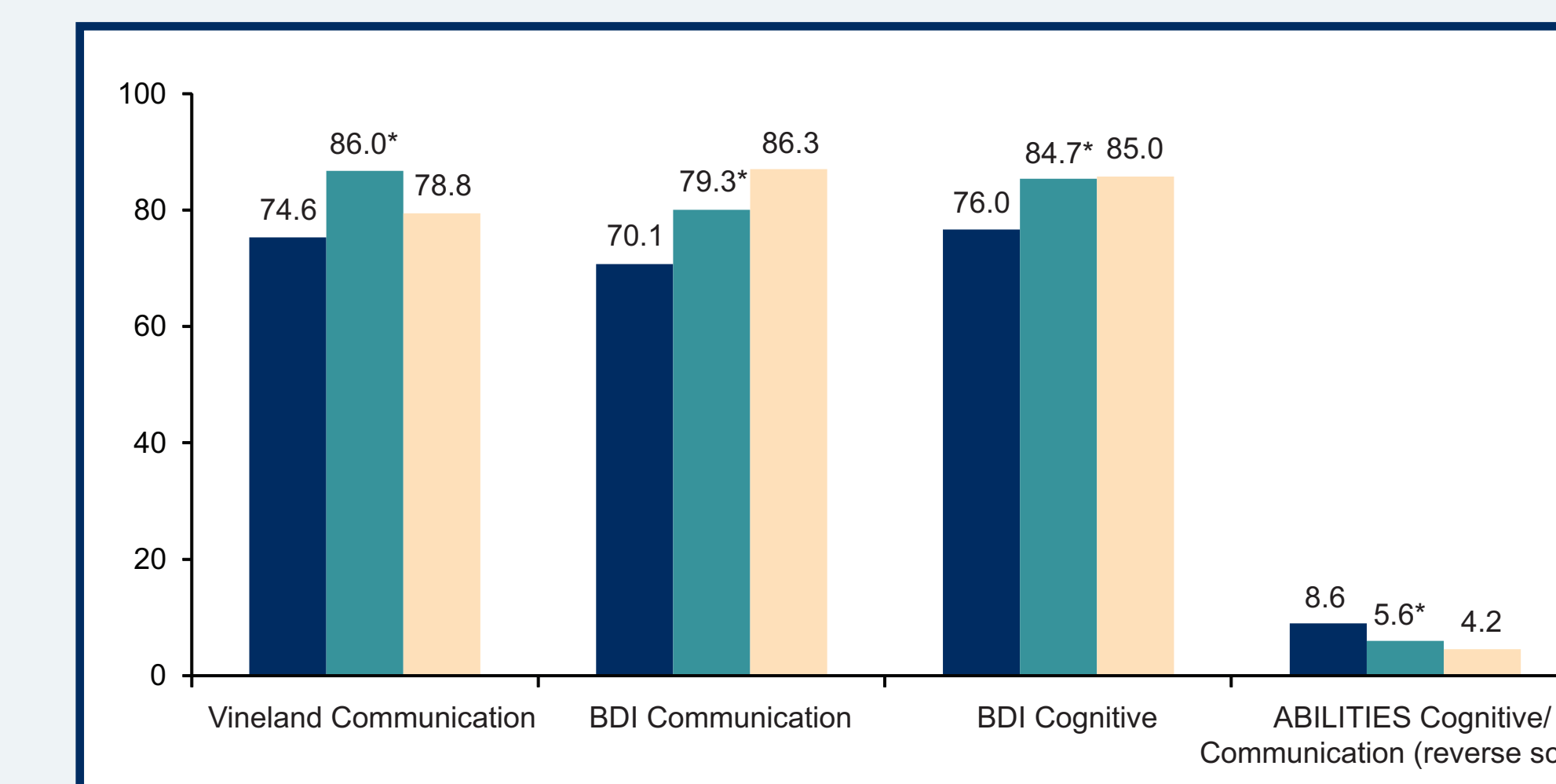
Findings

Figure 3. Positive social relationships: Mean assessment scores across COS groups: EI.



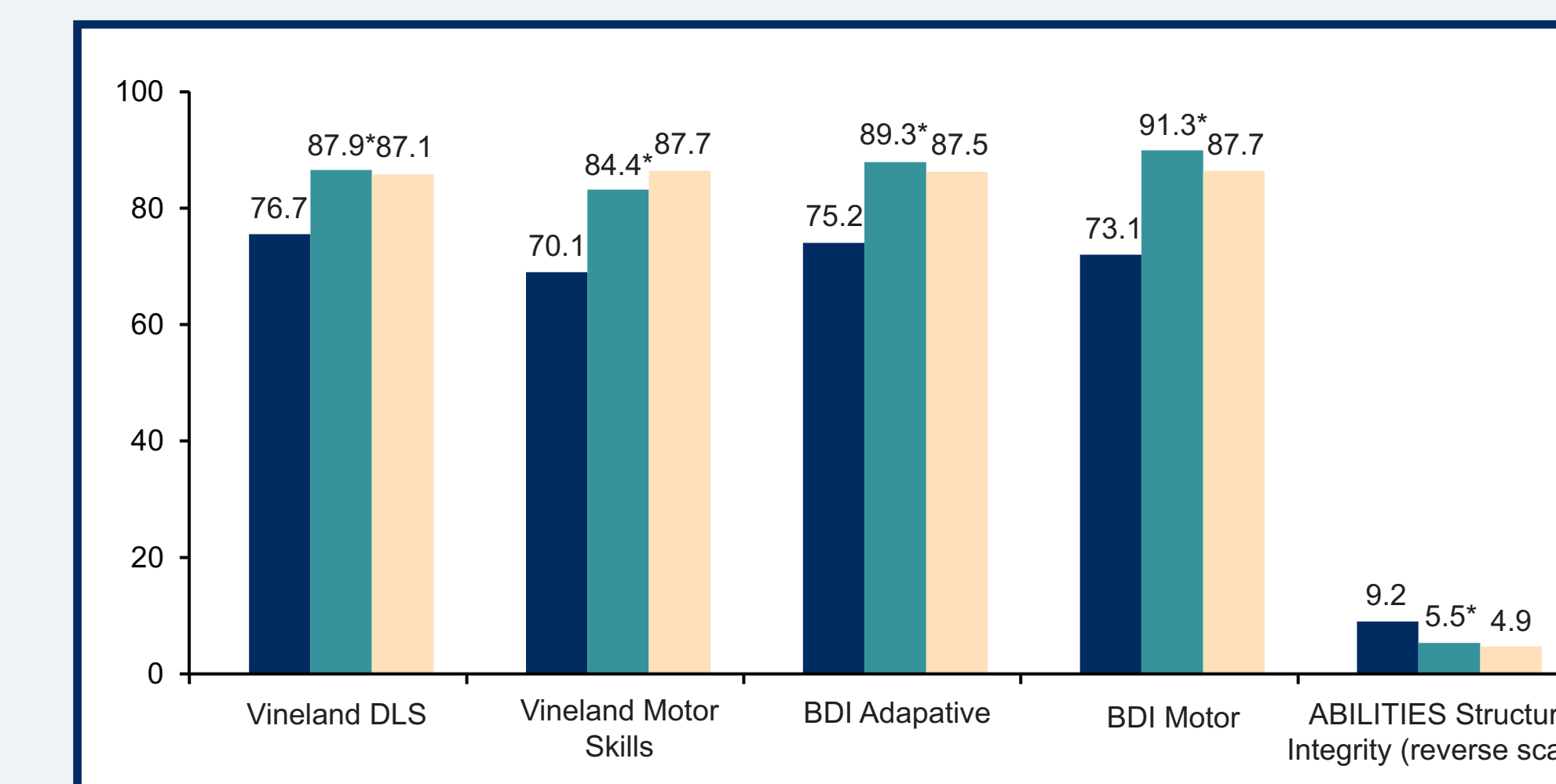
* p < .05.

Figure 5. Acquisition and use of knowledge and skills: Mean assessment scores across COS groups: EI.



* p < .05.

Figure 7. Use of appropriate behaviors to meet their needs: Mean assessment scores across COS groups: EI.



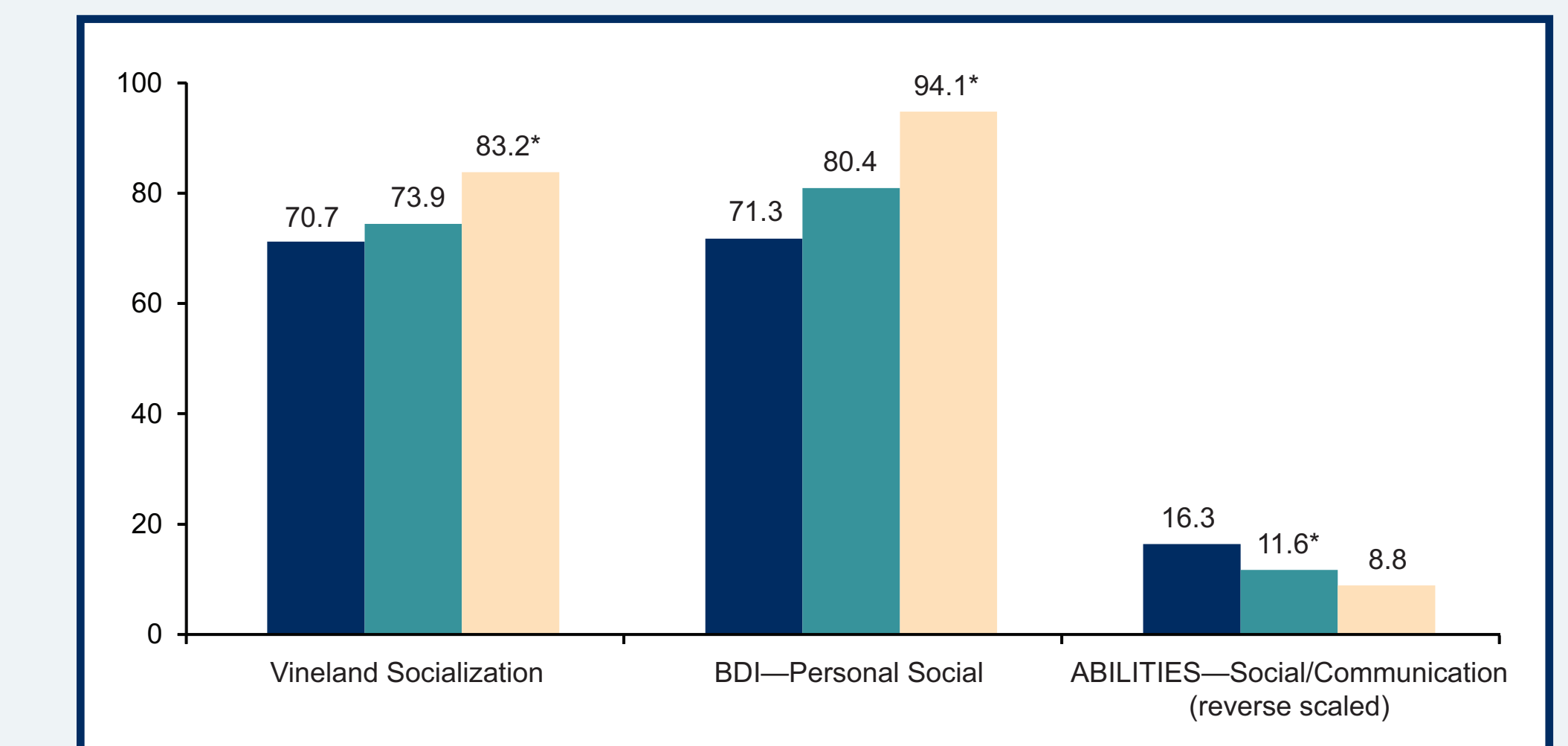
* p < .05.

Legend: COS Group 1 (dark blue), COS Group 2 (teal), COS Group 3 (orange)

Methodological Challenges

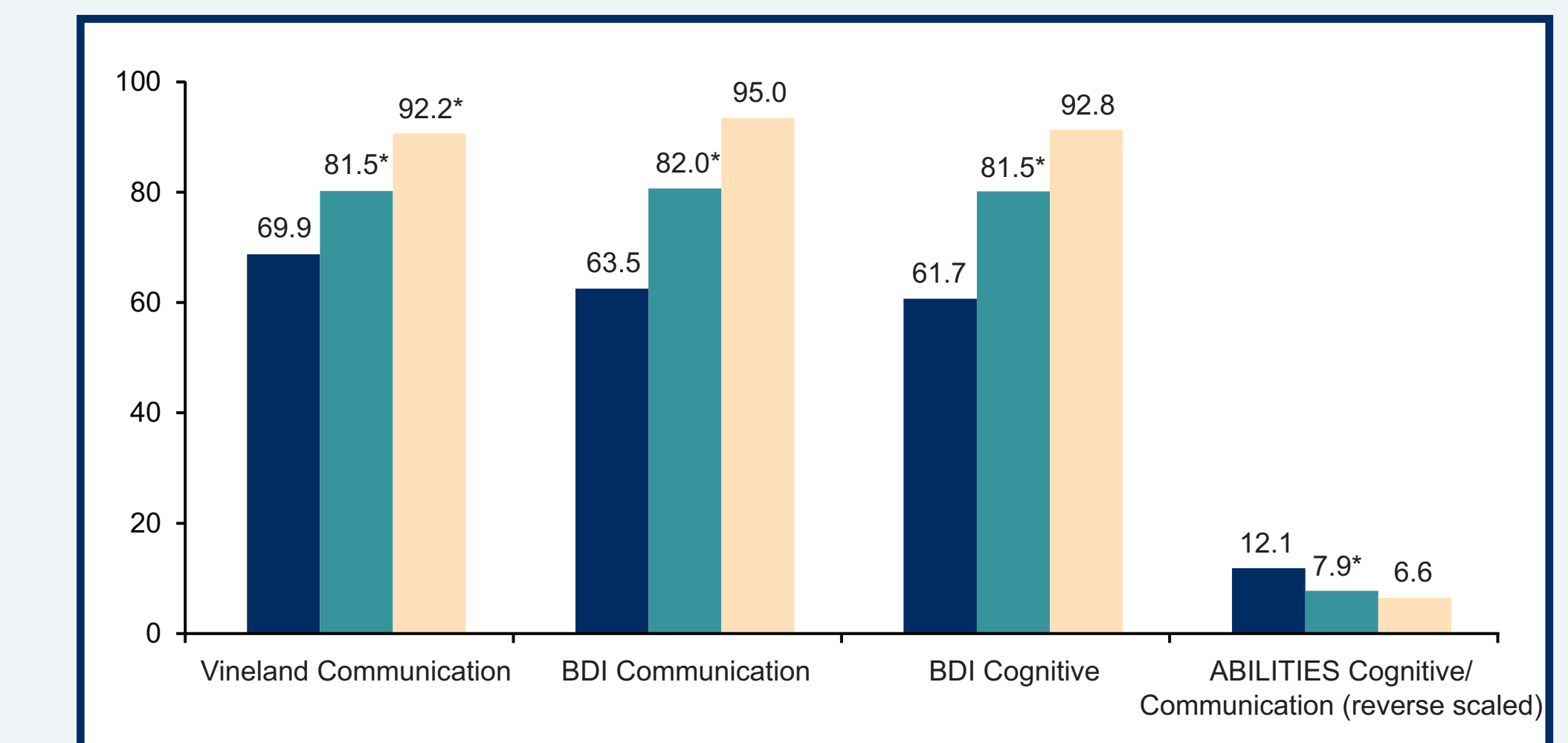
- Comparing outcomes to domains is an imperfect standard because functional outcomes, by definition, integrate skills across domains.
- Assessment tool subdomains provide content match, but they have poor reliability.
- COS ratings are not on an interval scale, a characteristic which limits the usefulness of correlations as evidence for validity claims.
- There is variation in COS implementation across programs and states, including the team process.
- Preliminary data suggest examining EI and ECSE separately may be important, but there was a limited sample size in preliminary data.
- Collapsing the seven COS rating points into three groups alters the metric used for accountability reporting.

Figure 4. Positive social relationships: Mean assessment scores across COS groups: ECSE.



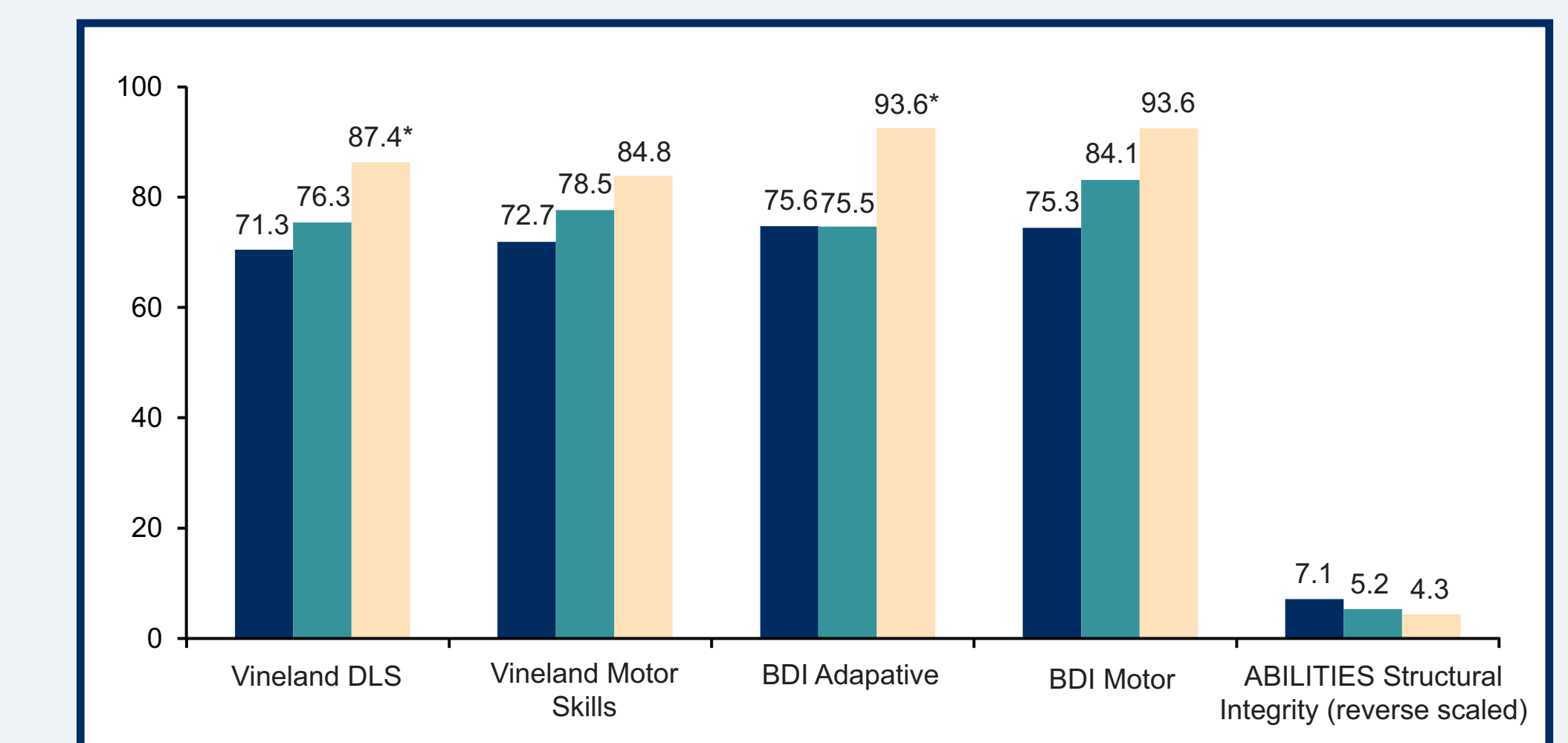
* p < .05.

Figure 6. Acquisition and use of knowledge and skills: Mean assessment scores across COS groups: ECSE.



* p < .05.

Figure 8. Use of appropriate behaviors to meet their needs: Mean assessment scores across COS groups: ECSE.



* p < .05.

Preliminary Results and Next Steps

- Overall, the patterns followed expected directions, showing:
 - Higher mean scores for higher COS groups on the expected subdomains of the Vineland-II and BDI-2.
 - Lower scores for higher COS groups on the ABILITIES Index.
- The overall pattern of the mean assessment scores for EI and ECSE supports the validity of the COS, but the results are preliminary.
 - Children in COS Group 1 had significantly lower scores on assessments than children in COS Group 3 across assessments, domains, and outcomes (with two exceptions for EI).
- Under Acquiring and Using Knowledge and Skills outcome, COS Group 1 scores were significantly lower than COS Group 2, across assessments and domains.
- Effect sizes were consistently larger for ECSE than EI. One exception: Larger EI than ECSE effect sizes found comparing motor domains to Taking Appropriate Action to Meet Needs.
- The relationship between the BDI and Vineland administered by the same assessor also was weaker in EI than ECSE.
- More data are needed to make final conclusions. Data continue to be collected.
- We hope to examine larger state datasets that have information about concurrent validity between assessment scores and COS ratings.