

# 7

## Technical Report

The SEAM was developed to address the need for psychometrically sound social-emotional assessment tools for young children. The SEAM was designed as a curriculum-based assessment measure to assist in the prevention and early identification of social-emotional difficulties and behavior disorders, as well as to build positive partnerships with families and optimize positive caregiver–child interactions in the first years of life.

SEAM benchmarks and items were identified from the literature on social-emotional development of young children raised in mainly Western cultures; certain concepts repeatedly emerged as those that were deemed essential or critically important to the mental health competence of young children (Squires & Bricker, 2007). These benchmarks and items were reviewed and revised in an iterative process based on feedback from family members and experts in infant mental health, early childhood, early intervention/early childhood special education, psychology, and behavior disorders. Any items that appeared difficult to understand or with ambiguous meanings were revised based on expert and caregiver feedback.

The psychometric properties of the SEAM were investigated in a series of research studies that are reported in this chapter. Psychometric studies on the Infant and Toddler Intervals were conducted as part of a federally funded research grant; pencil-and-paper as well as online data were gathered from a variety of caregivers served in programs around the United States. Additional data were also collected on Infant, Toddler, and Preschool Intervals through an online research web site (Squires et al., 2012a, 2012b). Research questions included the following:

- What is the item functioning for the Infant and Toddler Intervals?
- What is the reliability of the Infant, Toddler, and Preschool Intervals, including internal consistency, test–retest, and interrater reliability?
- What is the validity of the Infant, Toddler, and Preschool Intervals, specifically content and congruent validity?
- What is the utility of the SEAM system as rated by caregivers and early interventionists?

## SAMPLE

The pencil-and-paper data were gathered from caregivers and practitioners in early childhood programs serving typically developing children and children with developmental delays. Online data were gathered from a variety of caregivers around the United States using a research web site. Paper-and-pencil data, including demographic and utility surveys, were completed one of several ways, including individually by caregivers without practitioner assistance, by caregivers during an interview with a practitioner, and by a practitioner with at least 20 hours of weekly contact with a child (for interrater reliability). Practitioners included 1) early childhood classroom teachers and assistants primarily working with infants and toddlers who were typically developing and 2) early interventionists/early childhood special educators working with families and their children who were at risk or eligible for Part C services. Online measures, including demographic and utility surveys as well as SEAM intervals, were independently completed by caregivers, for the most part without assistance.

Data were collected in 49 states across the United States and from Canada. The number of completed SEAMs from each state ranged from 1 to 279, with the largest number coming from Oregon. The sample included a total of 2,201 SEAMs; 1,850 were collected online, and 351 were collected from paper-and-pencil versions. Of the sample, 59% of children were male and 41% were female. The children represented in the sample were predominately Caucasian (76.1%). Other ethnicities included multiracial (6.2%), Hispanic/Latino (4.9%), African American (4.7%), Asian (3.7%), American Indian/Alaskan Native (1.1%), Native Hawaiian/Pacific Islander (0.1%), other nonspecified ethnicity (1.4%), and unknown (1.8%). Fifty-eight percent of children were typically developing, whereas 42% of children were identified with a disability or developmental delay.

Data on family income and education level also were collected. The majority of caregivers reported incomes greater than \$50,000 (57%), whereas 43% reported incomes below that level. The greatest percentage of participating caregivers (60%) had a bachelor's or postgraduate/graduate degree, whereas 19% had some college, 17% had a high school diploma or general equivalency diploma, and 4% had not completed high school.

Data analysis techniques included item response theory (IRT) modeling as well as classical test analyses. IRT modeling was used to examine item order and fit statistics while traditional test analyses were employed to complete validity and reliability studies.

## PENCIL-AND-PAPER AND ONLINE DATA COLLECTION

Before proceeding to data analysis, a differential item functioning (DIF) analysis using a Rasch one parameter logistic (1PL) partial credit model (PCM) for polytomous scoring (Masters & Wright, 1997) was completed with the estimation software Winsteps 3.73 (Linacre, 2011) in order to examine whether SEAM items appeared to be functioning differently with different administration methods (i.e., paper/pencil and online) for the Infant and Toddler Intervals. The results from the DIF analysis indicated that there were only minor differences in item functioning between administration methods. Evidence for significant DIF was demonstrated in 3 out of 35 items (8.6%) in the Infant Interval and 2 out of 35 items (5.7%) in the Toddler Interval. These results suggested that most of the items functioned invariantly and were not affected by extraneous artifacts inherent in the method of completion by the rater. Similar results were found for ability status; 5 out of 35 items (14.3%) demonstrated DIF on the Infant Interval, whereas only 2 out of 35 items (5.7%) on the Toddler Interval had evidence of significant DIF. These results suggested minimal bias between groups. These findings supported the rationale for analyzing the data set as a whole.

## QUESTION 1: WHAT IS THE ITEM FUNCTIONING FOR THE INFANT AND TODDLER SEAM™ INTERVALS?

### Item Fit Statistics

Item fit statistics are generated as an indication of how well the selected model (i.e., Rasch 1PL PCM) fits the obtained data. Responses to items from people of varying estimated abilities should be consistent with the estimated item difficulty, such that participants with estimated high ability should be

**Table 7.1.** Item fit statistics

Age interval	Benchmark	Infit mean MNSQ (SD)	Infit MNSQ range	Outfit mean MNSQ (SD)	Outfit MNSQ range	Item ordering
Infant	1.0	0.98 (0.19)	0.84–1.29	0.95 (0.17)	0.81–1.23	<b>a, d, b, c</b>
	2.0	0.97 (0.21)	0.79–1.26	0.97 (0.24)	0.71–1.29	a, b, c
	3.0	0.99 (0.08)	0.89–1.09	0.97 (0.05)	0.89–1.02	b, a, c
	4.0	0.99 (0.10)	0.87–1.14	0.93 (0.14)	0.78–1.16	<b>b, c, d, a</b>
	5.0	1.00 (0.28)	0.74–1.44	1.16 (0.52)	0.74–2.01	b, a, c, d
	6.0	0.95 (0.16)	0.73–1.17	1.14 (0.31)	0.79–1.45	b, a, c, d
	7.0	1.00 (0.12)	0.87–1.16	0.95 (0.11)	0.84–1.10	<b>c, a, b</b>
	8.0	0.98 (0.09)	0.86–1.08	1.01 (0.14)	0.82–1.11	a, c, b
	9.0	1.00 (0.06)	0.92–1.06	0.92 (0.04)	0.87–0.97	a, b, c
	10.0	1.00 (0.11)	0.87–1.16	0.94 (0.15)	0.73–1.11	<b>d, a, c, b</b>
Toddler	1.0	1.00 (0.22)	0.73–1.23	1.00 (0.23)	0.74–1.25	<b>d, b, a, c</b>
	2.0	1.02 (0.26)	0.68–1.35	2.70 (3.12)	0.69–8.09	a, b, c, d
	3.0	0.99 (0.13)	0.84–1.17	1.03 (0.17)	0.84–1.25	a, b, c
	4.0	0.97 (0.12)	0.85–1.14	0.99 (0.12)	0.89–1.16	a, b, c
	5.0	0.99 (0.05)	0.92–1.06	1.00 (0.07)	0.92–1.08	b, a, c, e, d
	6.0	1.00 (0.06)	0.95–1.08	0.98 (0.04)	0.95–1.04	<b>c, b, a</b>
	7.0	0.98 (0.13)	0.83–1.14	0.98 (0.16)	0.79–1.19	a, c, b
	8.0	0.99 (0.11)	0.88–1.15	0.98 (0.12)	0.85–1.15	<b>c, d, a, b</b>
	9.0	0.98 (0.05)	0.94–1.03	0.93 (0.02)	0.92–0.95	a, b
	10.0	1.00 (0.10)	0.87–1.15	0.99 (0.12)	0.85–1.17	a, c, b, d
Preschool	1.0	0.99 (0.14)	0.76–1.17	0.92 (0.18)	0.65–1.20	<b>a, b, c, e, d</b>
	2.0	0.98 (0.14)	0.85–1.20	1.25 (0.43)	0.84–1.88	a, b, c, d
	3.0	0.99 (0.13)	0.83–1.20	0.97 (0.15)	0.79–1.15	<b>c, a, b, d</b>
	4.0	0.98 (0.04)	0.94–1.02	0.98 (0.04)	0.93–1.02	a, b
	5.0	0.99 (0.14)	0.80–1.18	1.00 (0.16)	0.80–1.24	<b>a, b, d, c</b>
	6.0	0.99 (0.15)	0.85–1.25	1.00 (0.21)	0.80–1.36	<b>a, d, b, c</b>
	7.0	0.98 (0.21)	0.80–1.28	0.94 (0.21)	0.78–1.24	<b>b, c, a</b>
	8.0	0.99 (0.13)	0.83–1.16	0.99 (0.16)	0.80–1.23	<b>e, b, a, c, d</b>
	9.0	0.99 (0.05)	0.93–1.05	0.93 (0.05)	0.87–0.99	<b>b, a, c</b>
	10.0	1.00 (0.12)	0.85–1.22	0.96 (0.14)	0.77–1.21	<b>g, e, f, a, b, d, c</b>

Letters are used instead of numbers for item ordering to facilitate visual analysis of the results (e.g., a = 1.1, b = 1.2, c = 1.3, d = 1.4, e = 1.5, f = 1.6, g = 1.7 for Preschool Benchmark 1.0). Italicized and bolded value indicates a misfit detected.

Key: MNSQ, mean square; SD, standard deviation.

able to demonstrate more difficult skills, whereas participants with lower ability should only be able to do easier items. Items that fit the model well are assigned fit statistics that range in value from 0.5 to 1.5. Items less than 0.5 are considered overly predictive, whereas items that are greater than 1.5 contain more noise than useful information and are considered degrading to the measure (Linacre, 2011). Confirming adequate model fit is a necessary step for ensuring credibility of results when performing an IRT modeling analysis. We examined item fit (i.e., outfit mean square) within each SEAM benchmark for this analysis. Results indicated that item-level fit statistics were well within the acceptable range for the majority of benchmarks, except for Item 5.1 (fit statistics = 2.01) from Benchmark 5.0 in the Infant Interval, Item 2.1 (fit statistics = 8.09) from Benchmark 2.0 in the Toddler Interval, and Item 2.1 (fit statistics = 1.88) from Benchmark 2.0 in the Preschool Interval (see Table 7.1). These results provide evidence of unidimensionality for each benchmark and support the use of the Rasch 1PL PCM as a means to evaluate item functioning.

## Item Functioning

Item functioning was evaluated in order to better understand the contribution of individual items within each benchmark of the SEAM. As previously mentioned, IRT offers a range of latent trait measurement models for explaining the relation between item responses and two classes of unobserved variables: 1) person ability and 2) item characteristics (Embretson & Reise, 2000; Hambleton & Swaminathan, 1985). Item characteristics (e.g., difficulty, sensitivity) are estimated with the person's responses to the set of measurement items, and each person's ability level is estimated based on his or her set of responses and the estimated item characteristics.

One of the purposes for doing the IRT modeling analysis was to examine the ordering of the items within each benchmark. Items within benchmarks on the experimental version of the SEAM

**Table 7.2.** Item functioning and Social-Emotional Assessment/Evaluation Measure (SEAM™) benchmarks with item order changes

Age interval	Benchmark	Original order and wording	Order from IRT modeling analysis	Final order	
Infant	C-1.0 Baby participates in healthy interactions.	1.1 Baby shows interest in you and other familiar caregivers.	1.1	1.1	
		1.2 Baby responds to you and other familiar caregivers.	1.4	1.3	
		1.3 Baby initiates and responds to communications.	1.2	1.4	
		1.4 Baby lets you know if she needs help or comfort.	1.3	1.2	
	C-4.0 Baby begins to show empathy for others.	4.1 Baby mimics your facial expressions.	4.2	4.1	
		4.2 Baby looks at and notices you and other familiar caregivers.	4.3	4.2	
		4.3 Baby looks at and notices others' emotional responses.	4.4	4.3	
		4.4 Baby responds to another's distress, seeking comfort for self.	4.1	4.4	
	C-7.0 Baby displays a positive self-image.	7.1 Baby laughs at, or smiles at, her image or picture of self.	7.3	7.2	
		7.2 Baby recognizes his name.	7.1	7.3	
		7.3 Baby calls attention to herself.	7.2	7.1	
	C-10.0 Baby shows a range of adaptive skills.	10.1 Baby eats and gains weight on schedule.	10.4	10.2	
		10.2 Baby eats a variety of age-appropriate foods.	10.1	10.4	
		10.3 Baby sleeps with few problems.	10.3	10.3	
		10.4 Baby eliminates (pees and poops) on regular schedule.	10.2	10.1	
	Toddler	C-1.0 Participates in healthy interactions	1.1 Toddler talks and plays with people whom she knows well.	1.4	1.3
1.2 Toddler initiates and responds to affection.			1.2	1.2	
1.3 Toddler initiates and responds when you communicate with her.			1.1	1.4	
1.4 Toddler lets you know if he needs help, attention, or comfort.			1.3	1.1	
C-6.0 Demonstrates independence		6.1 Toddler explores new environments, while maintaining some contact.	6.3	6.3	
		6.2 Toddler can separate from you in familiar environment with minimal distress.	6.2	6.2	
		6.3 Toddler tries new tasks before seeking help.	6.1	6.1	
C-8.0 Regulates attention and activity level		8.1 Toddler stays with motor activities for 5 minutes or longer.	8.3	8.3	
		8.2 Toddler looks at book or listens to a story for 5 minutes or longer.	8.4	8.4	
		8.3 Toddler moves from one activity to another without problems.	8.1	8.1	
		8.4 Toddler participates in simple games.	8.2	8.2	
Pre-school		C-3.0 Regulates social emotional responses	3.1 Child responds to peer's or caregiver's soothing when upset.	3.3	3.2
			3.2 Child can calm self when upset within 5 minutes.	3.1	3.3
			3.3 Child can calm self after periods of exciting activity.	3.2	3.1
			3.4 Child remains calm in disappointing situations.	3.4	3.4
		C-6.0 Demonstrates independence	6.1 Child explores new materials and settings.	6.1	6.1
	6.2 Child tries new task before seeking help.		6.4	6.3	
	6.3 Child stays with or returns to challenging activities.		6.2	6.4	
	6.4 Child can leave you without distress.		6.3	6.2	
	C-7.0 Displays positive self-image	7.1 Child knows personal information.	7.2	7.3	
		7.2 Child shows off work, takes pride in accomplishments.	7.3	7.1	
		7.3 Child makes positive statements about self.	7.1	7.2	
	C-8.0 Regulates attention and activity level	8.1 Child stays with motor activity for 10 minutes or longer.	8.5	8.3	
		8.2 Child participates in early literacy activities.	8.2	8.2	
		8.3 Child moves from one activity to another without problems.	8.1	8.4	
		8.4 Child participates in games with others.	8.3	8.5	
		8.5 Child regulates his activity level to match setting.	8.4	8.1	
	C-10.0 Shows a range of adaptive skills	10.1 Child feeds self and eats a variety of foods without a problem.	10.7	10.4	
		10.2 Child dresses self.	10.5	10.5	
		10.3 Child goes to bed and falls asleep without a problem.	10.6	10.7	
		10.4 Child uses the toilet appropriately.	10.1	10.6	
		10.5 Child manages changes in settings and conditions.	10.2	10.2	
		10.6 Child keeps himself safe in potentially dangerous conditions.	10.4	10.3	
		10.7 Child solves problems to meet her needs.	10.3	10.1	

Key: IRT, item response theory.

were intended to be ordered from easier to more difficult to facilitate an examination of a child's progress in social-emotional competencies. This initial item ordering was based on developmental quotient, or the relative difficulty of skills from the literature on social-emotional development of young children. Results from the estimated item difficulty suggested that the majority of items within each SEAM benchmark were in fact hierarchically organized (i.e., the numerical order reflects the developmental hierarchy of items, with "x.1" as the easiest) and confirmed the predetermined developmental hierarchy of these social-emotional skills (see Table 7.1). Benchmarks remained in the predetermined item order when only minor item order changes were indicated (e.g., switching the order of two adjacent items). The item order within four benchmarks on the Infant Interval, three benchmarks in the Toddler Interval, and five benchmarks in the Preschool Interval was found to be largely different from the hypothesized order (see Table 7.2). According to the IRT modeling analysis of item responses, the preidentified easier or easiest item within these benchmarks was, in fact, more difficult or in some cases appeared to be the most difficult item. Researchers considered the IRT suggested ordering of items for the benchmarks in which disorder was detected and carefully examined each set of items to determine whether the disagreement appeared to be due to 1) true item misorder or 2) misinterpretation of the item by respondents. The decision was then made to either reorder the items as suggested by the IRT results or to keep the items in their original position. Items were reordered according to the IRT results in most cases; however, original item ordering was maintained for a few benchmarks, and minor editing was done (either to the item itself, its accompanying example[s], or both) in an attempt to increase the clarity of individual items.

## QUESTION 2: WHAT IS THE RELIABILITY OF THE INFANT, TODDLER, AND PRESCHOOL SEAM™ INTERVALS, INCLUDING INTERNAL CONSISTENCY, TEST–RETEST, AND INTER-RATER RELIABILITY?

### Interrater Reliability

Interrater reliability data for the Infant and Toddler Intervals were collected from teacher dyads working at a high-quality child care center serving primarily children of University of Oregon faculty and staff. Master teachers and assistant teachers from the infant and toddler classrooms participated. Pearson product moment correlation coefficients and intraclass correlations were computed to examine interrater agreement. Results are presented in Table 7.3 for four teacher dyads (one dyad for the Infant Interval of the SEAM,  $n = 12$  children) and three dyads for the Toddler Interval of the SEAM (Toddler Class 1,  $n = 7$  children; Toddler Class 2,  $n = 7$  children; and Toddler Class 3,  $n = 8$  children). The Pearson product moment correlation coefficient ( $r = .776$ ) was significant at  $p < .01$  for the Infant Interval of the SEAM. The Pearson product moment correlation coefficient for Toddler Class 2 ( $r = .948$ ) was also significant at  $p < .01$ . Pearson product moment correlation coefficients for Toddler Classes 1 and 3 were not significant. Intraclass correlations were also analyzed to examine the consistency of differences between scores across raters. Results of the intraclass correlations were strong and significant for teachers in the infant classroom and for toddler teachers in Classes 1 and 2, but were not significant for the teachers in Toddler Class 3 (see Table 7.3).

### Test–Retest Reliability

Test–retest reliability data were collected by online caregiver participants. After completing the SEAM via a research web site, caregivers were immediately given the option to complete a second SEAM,

**Table 7.3.** Correlations of total Social-Emotional Assessment/Evaluation Measure (SEAM™) scores between professional raters within classrooms

Classroom	<i>n</i>	<i>r</i>	Intraclass correlation
Infant 1	12	.776**	.564*
Toddler 1	7	.668	.657*
Toddler 2	7	.948**	.932**
Toddler 3	8	.640	.324

Key: \* $p < .05$ ; \*\* $p < .01$ .

**Table 7.4.** Test–retest correlations

SEAM™ interval	<i>n</i>	<i>r</i>
Infant	43	.987**
Toddler	42	.968**
Preschool	49	.989**

Key: \**p* < .05; \*\**p* < .01; SEAM™; Social-Emotional Assessment/Evaluation Measure.

**Table 7.5.** Correlations between Infant Social-Emotional Assessment/Evaluation Measure (SEAM™) benchmarks and overall SEAM scores

Benchmark	1	2	3	4	5	6	7	8	9	Total score
1.0										.85
2.0	.65									.83
3.0	.40	.36								.73
4.0	.66	.59	.42							.84
5.0	.62	.56	.36	.67						.85
6.0	.41	.56	.28	.47	.53					.82
7.0	.58	.59	.32	.62	.67	.65				.83
8.0	.55	.55	.35	.58	.64	.63	.67			.83
9.0	.36	.31	.37	.37	.36	.30	.31	.42		.69
10.0	.35	.31	.49	.36	.41	.44	.43	.44	.41	.77

Note: All correlations are significant at *p* < .01. Total number of Infant SEAMs included in the analyses between benchmarks ranged from 1,130 to 1,134 and was 1,153 for benchmark correlations with SEAM total score.

**Table 7.6.** Correlations between Toddler Social-Emotional Assessment/Evaluation Measure (SEAM™) benchmarks and overall SEAM scores

Benchmark	1	2	3	4	5	6	7	8	9	Total score
1.0										.87
2.0	.54									.84
3.0	.35	.32								.78
4.0	.56	.66	.33							.82
5.0	.62	.43	.41	.58						.90
6.0	.42	.36	.43	.44	.63					.82
7.0	.59	.65	.31	.68	.56	.43				.79
8.0	.49	.45	.51	.49	.62	.58	.55			.88
9.0	.51	.39	.42	.48	.62	.49	.58	.59		.84
10.0	.45	.46	.49	.50	.47	.47	.53	.51	.53	.83

Note: All correlations are significant at *p* < .01. Total number of Toddler SEAMs included in the analyses between benchmarks ranged from 467 to 472 and was 490 for benchmark correlations with SEAM total score.

blind to the results of the first one. Results indicated strong, significant agreement between the two SEAM completions for all three intervals (see Table 7.4).

## Internal Consistency

Internal consistency of the SEAM was addressed by examining the relation between average benchmark scores using correlational analyses and Cronbach's coefficient alpha (Cronbach, 1951). Pearson product moment correlation coefficients between benchmarks ranged from .28 to .67 for the Infant Interval of the SEAM, from .31 to .68 for the Toddler Interval of the SEAM, and from .41 to .81 for the Preschool Interval of the SEAM (see Tables 7.5, 7.6, and 7.7). In addition, the correlational analyses between benchmarks and overall SEAM scores were consistent, ranging from .69 to .85 for the Infant Interval, from .78 to .90 for the Toddler Interval, and from .73 to .88 for the Preschool Interval. All correlations were significant, suggesting congruence between benchmarks within each age interval as well as between benchmarks and total SEAM scores. Cronbach's coefficient alphas were also calculated for each age interval. The standardized alpha was .90 for the Infant Interval, .91 for the Toddler Interval, and .96 for the Preschool Interval, indicating strong internal consistency.



**Table 7.7.** Correlations between Preschool Social-Emotional Assessment/Evaluation Measure (SEAM™) benchmarks and overall SEAM scores

Benchmark	1	2	3	4	5	6	7	8	9	Total score
1.0										.88
2.0	.76									.78
3.0	.60	.47								.74
4.0	.74	.69	.56							.78
5.0	.81	.67	.57	.70						.87
6.0	.61	.51	.50	.51	.58					.73
7.0	.69	.67	.41	.55	.65	.52				.76
8.0	.73	.61	.64	.64	.75	.60	.64			.88
9.0	.62	.49	.63	.54	.62	.47	.52	.71		.78
10.0	.61	.55	.61	.52	.60	.59	.55	.73	.66	.84

Note: All correlations are significant at  $p < .01$ . Total number of Preschool SEAMs included in the analyses between benchmarks ranged from 604 to 653 and was 524 for benchmark correlations with SEAM total score.

**Table 7.8.** Mean Social-Emotional Assessment/Evaluation Measure (SEAM™) scores and correlations with age across 6-month intervals

Age in months	<i>n</i>	<i>M</i>	<i>r</i>
SEAM for Infants	235		.354**
0–6	49	84.32	
6–12	153	93.49	
12–18	33	97.61	
SEAM for Toddlers	56		.391**
18–24	20	82.39	
24–30	19	86.83	
30–36	17	92.59	
SEAM for Preschool	240		.124
36–42	79	103.82	
42–48	62	107.89	
48–54	53	106.42	
54–60	35	108.49	
60–66	11	106.91	

Key: \* $p < .05$ ; \*\* $p < .01$ .

Note: Age ranges begin on the first day of the first month indicated and end the day before the last month indicated. For example, the 6–12 month age range includes children who are between 6 months, 0 days old and 11 months and 30/31 days old.

### QUESTION 3: WHAT IS THE VALIDITY OF THE INFANT, TODDLER, AND PRESCHOOL SEAM™ INTERVALS, SPECIFICALLY CONTENT AND CONGRUENT VALIDITY?

#### Correlation of Mean SEAM™ Scores with Age

Two analyses were computed using a subset of the data sample that included children who were known to be typically developing in order to 1) calculate mean SEAM scores across 6-month intervals for all age intervals and 2) calculate correlation of mean SEAM scores with age for the Infant and Toddler Intervals. There was a consistent increase in mean scores across the 6-month age intervals in both the Infant and Toddler Intervals (see Table 7.8). Nonetheless, the Preschool Interval did not demonstrate this increasing trend. Correlations of mean scores with age for the Infant ( $r = .354$ ) and Toddler ( $r = .391$ ) Intervals were moderate and significant at  $p < .01$ , suggesting that children's scores did increase with age but with some variations, which means children of the same age may have different total scores on the SEAM. Correlation of mean scores with age for the Preschool Interval ( $r = .124$ ) was low and not significant. Lack of significance for children of preschool age might be due to the variability of children's social-emotional skills related to whether they attend preschools or other learning centers and the skills they learn in these settings. Many skills in the social-emotional domain are acquired by infants and toddlers based on developmental maturation rather than experiential learning. SEAM authors also expected that qualitative, rather than quantitative, changes might be observed at the preschool level because children have generally acquired basic social-emotional skills by this age. Age-based differences are observed in the quality and sophistication of their demonstration of these acquired skills.

**Table 7.9.** Correlations between Social-Emotional Assessment/Evaluation Measure (SEAM™) scores and other related measures

		DECA	ITSEA Compliance	ITSEA Negative Emotion	ITSEA Pro-social	ASQ:SE
Infant SEAM scores	<i>r</i>	.754**	.628**	-.415*	.651**	-.557**
	<i>n</i>	13	27	26	24	860
Toddler SEAM scores	<i>r</i>	NA	.564**	-.261**	.652**	-.516**
	<i>n</i>		119	120	120	162
Preschool SEAM scores	<i>r</i>	NA	NA	NA	NA	-.810**
	<i>n</i>					417

Key: \* $p < .05$ ; \*\* $p < .01$ ; ASQ:SE, Ages & Stages Questionnaires®: Social-Emotional (Squires, Bricker, & Twombly, 2002); DECA, Devereux Early Childhood Assessment Infant-Toddler (Mackrain, LeBuffe, & Powell, 2007); ITSEA, Infant Toddler Social Emotional Assessment (Carter & Briggs-Gowan, 2006).

### Concurrent Validity

The Devereux Early Childhood Assessment Infant-Toddler (DECA-IT; Mackrain, LeBuffe, & Powell, 2007), ITSEA (Carter & Briggs-Gowan, 2006), and ASQ:SE (Squires, Bricker, & Twombly, 2002) were used as the criterion measures to examine the concurrent validity of the SEAM. The Infant Interval of the SEAM was compared with the DECA-IT, ITSEA, and ASQ:SE, and the Toddler Interval of the SEAM had ITSEA and ASQ:SE as the criterion measures.

**Devereux Early Childhood Assessment Infant-Toddler** Correlations for DECA-IT scores with the Infant Interval ( $n = 13$ ) were strong and significant ( $r = .754$ ). Results are shown in Table 7.9.

**Infant Toddler Social Emotional Assessment** Correlations were strong and significant for both the Infant ( $n = 27$ ) and Toddler ( $n = 120$ ) Intervals for both Compliance and Pro-social domains (see Table 7.9). The number of subjects varied slightly across domains because cases were included only if all items within a domain were scored, thus allowing a total domain score to be calculated. The correlation between Infant Interval scores and the Compliance domain was  $r = .628$ , and the Toddler Interval was  $r = .564$ . The correlation with the Pro-social domain was  $r = .651$  for the Infant Interval and  $r = .652$  for the Toddler Interval. As expected, correlations between the Negative Emotion subscale for both Infant and Toddler Intervals were in a negative direction. Although the Infant Interval results were strong and significant ( $r = -.415$ ), the Toddler Interval results showed a weak correlation ( $r = -.261$ ) with the Negative Emotion subscale.

**Ages & Stages Questionnaires®: Social-Emotional** Correlations with the ASQ:SE were in a negative direction for both Infant ( $n = 860$ ) and Toddler ( $n = 162$ ) Intervals for the total sample (measuring challenging behaviors in ASQ:SE and competence in SEAM) (see Table 7.9). The correlation between the ASQ:SE and the Infant Interval ( $r = -.557$ ) was strong and significant, and the correlation between the ASQ:SE and the Toddler Interval ( $r = -.516$ ) was moderate/strong and also significant. The correlation between the ASQ:SE and the Preschool ( $n = 417$ ) Interval ( $r = -.810$ ) was also strong and significant. This was the expected outcome because SEAM scores increased with competence and ASQ:SE scores increased as negative behaviors and concerns increased.

### QUESTION 4: WHAT IS THE UTILITY OF THE SEAM™ SYSTEM?

The authors collected utility data from 434 caregivers who completed the SEAM. More than 93% of caregivers felt that the SEAM asked appropriate and useful questions. Ninety-one percent of caregivers felt that items were clearly worded. Caregivers indicated they were alerted to new child skills (56% agreed or strongly agreed, whereas 23% had no opinion); 89% indicated that completing the SEAM did not bring up any concerns about their children that they felt they needed to talk to someone about. Caregivers said that it took them an average of 9 minutes to complete the SEAM, indicating a reasonable time for caregiver completion.



Researchers also conducted a written utility survey with 35 practitioners from Part C early intervention programs. Of this group, demographic information was collected from 34 practitioners. Practitioners had an average of 8 years of experience working with children birth through age 5 years. The majority held either a bachelor's (47%) or postgraduate/graduate degree (47%), with 6% holding an associate's degree. Practitioners used a 4-point scale to rate their skill level related to providing mental health services to infants and toddlers and their families, with 1 = *very low* and 4 = *very high*. Four participants (12%) gave themselves a 1 rating; 19 participants (56%) gave a 2 rating; 10 participants (29%) gave a 3 rating; and 1 participant (3%) gave a 4 rating.

Six percent of practitioners used only the Infant Interval of the SEAM; 47% used only the Toddler Interval; and 47% of practitioners used both. Each practitioner completed between 1 and 19 SEAM intervals, with the majority completing 1 to 4. The majority of practitioners (91%) completed the SEAM with families during home visits; whereas others (6%) completed the SEAM in a child care center or in other ways (11%), such as having a caregiver complete it on his or her own at home. Written comments on preferred completion methods revealed a preference for completing the SEAM with caregivers during home visits, through a conversational or interview style that permitted discussion of questions.

Ninety-two percent of practitioners agreed or strongly agreed that SEAM items were clear and easy to understand. Seventy-nine percent ( $n = 33$ ) agreed or strongly agreed that completing the SEAM gave them meaningful information about a child's social-emotional abilities and needs. Sixty percent ( $n = 33$ ) agreed or strongly agreed that they would use the SEAM again; 30% had no opinion; 9% disagreed. Sixty-two percent ( $n = 29$ ) agreed or strongly agreed that they planned to address some of the skills caregivers indicated as intervention goals on the SEAM; 35% had no opinion.

Researchers also conducted two focus groups in order to further understand the utility of the SEAM. The first focus group included 12 teachers and took place in an early child care setting following the interrater reliability study described in which the SEAM was completed by teachers. The second focus group was conducted with five practitioners (home visitors and toddler classroom teachers) from an agency responsible for delivering Part C early intervention services. Participating practitioners in this group completed both the SEAM and the SEAM Family Profile. These practitioners completed between 12 and 41 SEAM protocols each and reported that some caregivers independently completed the SEAM with no problems, while the practitioners administered the SEAM to others in an interview format or in the context of a guided parent group. Focus group questions asked about the benefits and challenges of using the SEAM system with families and included questions such as 1) Did using the SEAM affect your relationship with the families you serve? If so, how? and 2) Did the SEAM give you any new information about the children and families with whom you work?

Although a formal qualitative study was not completed, focus group questions were intended to solicit more in-depth information than the utility surveys provided. The data reported next are summarized from the group conducted with Part C providers who implemented the SEAM in a caregiver-completed format. During this focus group process, facilitators regularly checked for agreement or disagreement among participants. The themes summarized next are those on which participants expressed agreement or which were expressed by the majority of participating practitioners.

Focus group participants preferred the interview format and stressed the importance of having a practitioner involved in SEAM completion in order to clarify items, explain examples, explore concerns, and help caregivers choose appropriate response options. Other themes emerged surrounding the impact of the SEAM on the relationship between practitioners and families. Several practitioners indicated that the SEAM opened up conversations with families that might not otherwise have taken place. One practitioner stated,

*"So I think it was a really nice forum for a conversation to learn more about the frequency. Things were more intense than I had realized previously. I hadn't really asked the right questions until I stepped through it."*

Practitioners also indicated that they learned new information about families that helped guide future interventions and support children's development. According to one practitioner,

*“By doing [the SEAM] with [families], I found some areas that I didn’t know were issues for families or where they were having difficult times. So it did help me guide the home visit piece of my service. We never had a conversation about it because the social area wasn’t an area that we were too concerned about. It doesn’t mean it’s an area of delay, but it does mean it’s something that’s affecting the family that I can help with, which overall helps with overall development.”*

Finally, practitioners discussed the difficulty of using the SEAM with families whose children had serious concerns of which caregivers were not already aware. They stressed the importance of sensitivity to each family’s individual issues and needs and cautioned against having every family independently complete the SEAM, particularly when it might reveal new information that might be emotionally difficult to receive. One practitioner described one such family to whom she had given the SEAM.

*“It was a child who we were truly concerned about with autism, and the parent got very emotional because it really brought out the social differences for that family. I hadn’t really thought about how much this would affect that particular diagnosis or the ability to see where the discrepancies were. It definitely was a trigger. I would have done it really differently because it really hit them hard, and I felt like it was an emotional effect that I would have liked to have buffered had I really thought about it. And now if I had to do it again, I would do it differently, for sure, or maybe not even do it with that particular family.”*

## **FUTURE RESEARCH EFFORTS**

The authors have collected objective data in this initial study of the SEAM system to substantiate the SEAM intervals as well as additional utility data gathered through focus groups. The authors used IRT to investigate item functioning and item ordering and conducted classical test analyses in order to perform the validity and reliability studies. Results suggest the SEAM had robust results related to validity, reliability, and utility. Further study is needed, however, with a stratified, randomized national sample to confirm these results. Linked system processes with intervention/curriculum development and program evaluation, including child monitoring and program effectiveness, also need to be studied.

We believe the SEAM is a measure with data that validates its use in the realm of social-emotional assessment and intervention. Research is ongoing; the authors continue to collect and assemble information to examine the validity, reliability, and usefulness of the SEAM. Additional recent research efforts have begun to investigate the SEAM Family Profile as well as the quality of goals written by practitioners using both the SEAM and the SEAM Family Profile. Future research efforts will focus on the effectiveness of the SEAM in monitoring child progress over time.