Measuring the Quality of Caregiver-Child Interactions for Infants and Toddlers (Q-CCIIT)

Appendices

January 2015
OPRE Report 2015-13

Q-CCIIT

This page has been left blank for double-sided copying.

APPENDIX A

QUALITY OF CAREGIVER-CHILD INTERACTIONS FOR INFANTS AND TODDLERS (Q-CCIIT) FIELD TEST: CAREGIVER QUESTIONNAIRE

Appendix A includes a copy of the caregiver questionnaire which was used in the field test to gather information about caregivers' experience and the structural features of classrooms and family child care settings. These measures were used to explore the validity of the Q-CCIIT observation instrument (the caregiver questionnaire is not part of the Q-CCIIT observation instrument).

This page has been left blank for double-sided copying.

OMB No.: 0970-0392
Expiration Date: 09/30/2013

MATHEMATICA
Policy Research

Quality of Caregiver-Child Interactions for Infants and Toddlers (Q-CCIIT): Caregiver Questionnaire

Sept 17, 2012

FOR OFFICE USE ONLY: LABEL

ABOUT THIS QUESTIONNAIRE

This questionnaire is an important part of a larger study supported under a contract from the U.S. Department of Health and Human Services, Administration for Children and Families. The overall purpose of the Quality of Caregiver-Child Interactions for Infant and Toddlers (Q-CCIIT) project is to understand the ways caregivers interact with infants and toddlers in center-based and family child care. Participation in this project is voluntary.

This form requests information about your child-care setting and your background and experience. The information will be used for research purposes only and will be kept confidential to the extent allowed by law. Your answers to these questions will not be shared with your employer. Your name will not be attached to any information you give us. Please note that pages are double-sided, and please complete the entire 9 pages of the questionnaire, but you may skip any question you do not wish to answer.

Most of the questions can be answered by marking an "X" in the box. For a few questions you may be asked to write in a response.

1 ☐ 2 ☒ 3 ☐

Thank you very much for your help.

A. STAFF DEVELOPMENT

A1. Please record today's date:

|__|__| / |__|__| / |_2_|_0_|__|__|

A2. How many hours a year do you attend staff trainings?

|__|__|__| HOURS

N = 839
MEDIAN 20.0; MEAN 35.9
RANGE 0-950
MISSING 13.7%

A3. How often do you have one-on-one supervision meetings or group supervision meetings?

MARK ONE ONLY

7.0% 0 ☐ Never
5.3% 1 ☐ Once a year
16.8% 2 ☐ A few times a year
3.7% 3 ☐ Every 2 months
30.4% 4 ☐ Once a month
8.0% 5 ☐ Twice a month
12.2% 6 ☐ Once a week
6.0% 7 ☐ More than once a week
8.0% n/a ☐ Not applicable

Missing 2.5%

A4. Is there someone who mentors you in your classroom, that is, someone who observes your teaching on a regular basis and provides feedback, guidance, and training?

61.0% 1 ☐ Yes
36.5% 0 ☐ No

Missing 2.5%

A5. Are <u>you</u> a member of a professional support network such as the Family Day Care Professional Association or the National Association for the Education of Young Children (NAEYC)?

33.9% 1 ☐ Yes
62.4% 0 ☐ No → **GO TO A7**

Missing 3.8%

A6. If yes, do you meet on a regular basis with other caregivers as part of a support network?

19.2% 1 ☐ Yes
11.8% 0 ☐ No

62.4% Logical skip

Missing 6.6%

A7. Does your child care setting provide you with any of the following?

	YES	NO	DON'T KNOW	MISSING
a. Tuition reimbursement for relevant college courses.....	1 ☐ 39.1%	0 ☐ 42.4%	d ☐ 14.6%	**3.9%**
b. Reimbursement for workshop fees or other costs for outside training.....	1 ☐ 49.6%	0 ☐ 34.5%	d ☐ 13.0%	**3.0%**
c. Time during work hours for staff development activities such as attending courses or workshops	1 ☐ 60.7%	0 ☐ 28.8%	d ☐ 7.7%	**2.8%**

B. COMMUNICATION WITH PARENTS

B1. There are many different ways that program staff can share information with parents. Do you use any of the following to communicate with parents?

MARK ONE PER ROW

	YES	NO	Missing
a. Newsletters	1 ☐ 76.3%	0 ☐ 17.9%	**5.8%**
b. Daily logs	1 ☐ 79.7%	0 ☐ 14.6%	**5.7%**
c. Personal/individualized notes	1 ☐ 82.0%	0 ☐ 10.7%	**7.3%**
d. Email/internet/website	1 ☐ 56.9%	0 ☐ 32.0%	**11.2%**
e. Flyers	1 ☐ 70.4%	0 ☐ 19.8%	**9.9%**
f. Posted notices	1 ☐ 82.0%	0 ☐ 10.7%	**7.3%**
g. Verbal (in-person or by phone)	1 ☐ 95.1%	0 ☐ 2%	**3.1%**
h. Other *(Please specify)* _____	1 ☐ 7.4%	0 ☐ 31.6%	**61.0%**

B2. How often do you talk to parents about how their children are doing on a formal or informal basis?

MARK ONE ONLY

2.3% 0 ☐ Never

2.0% 1 ☐ Only at parent-teacher conferences

1.0% 2 ☐ Every 2 or 3 months

1.7% 3 ☐ Once or twice a month

9.2% 4 ☐ Once or twice a week

83.0% 5 ☐ Daily

Missing 0.9%

B3. How often do you hold formal parent-teacher conferences with parents about individual children?

MARK ONE ONLY

21.6% 0 ☐ Never

8.4% 1 ☐ Once a year

39.5% 2 ☐ Twice a year

9.8% 3 ☐ 3 times a year

16.7% 4 ☐ 4 or more time a year

Additional code based on hard copy comments

0.2% 5 ☐ As needed

Missing 3.6%

0.2% Not Applicable

C. EMPLOYMENT

C1. Are you currently working at your child care setting full or part-time?

MARK ONE ONLY

81.1% 1 ☐ Full-time

18.1% 0 ☐ Part-time

Missing 0.8%

C2. Counting this school year, how long have you worked in your <u>current child care setting</u>?

|__|__| YEARS |__|__| MONTHS

Missing 2.3%

N = 950
MEDIAN 4.0; MEAN 6.0
RANGE 0-39

C3. Counting this school year, how long have you worked in your <u>current classroom</u>?

|__|__| YEARS |__|__| MONTHS

Missing 4.8%

N = 925
MEDIAN 1.7; MEAN 3.8
RANGE 0-39

C4. How likely are you to continue working in any child care setting next year?

MARK ONE ONLY

7.6% 1 ☐ Very unlikely

2.3% 2 ☐ Somewhat unlikely

10.9% 3 ☐ Somewhat likely

78.4% 4 ☐ Very likely

Missing 0.8%

C5. Please indicate your role(s) at this child care setting.

MARK ALL THAT APPLY

12.5% 1 ☐ Owner

9.1% 2 ☐ Director

30.0% 3 ☐ Lead Teacher

31.6% 4 ☐ Assistant Teacher

32.2% 5 ☐ Teacher

3.5% 6 ☐ Administrative Assistant

3.5% 7 ☐ Other role (*please specify*) _____

Additional codes based on frequent "Other" specified

0.9% 8 ☐ Substitute

1.3% 9 ☐ Volunteer (includes parent and foster grandparent program)

1.3% 10 ☐ Floater

0.6% 11 ☐ Special teacher (i.e., music teacher)

Missing 0.3%

C6. Are you a parent?

58.4% 1 ☐ Yes

40.3% 0 ☐ No → **GO TO C9**

Missing 1.2%

C7. If yes, have any of your children been enrolled in the child care setting where you are employed?

24.8% 1 ☐ Yes

32.9% 0 ☐ No → **GO TO C9**

40.3% logical skip

Missing 2.0%

C8. If yes, are any of your children <u>currently</u> in your classroom?

3.6% 1 ☐ Yes

20.7% 0 ☐ No

73.2% logical skip

Missing 2.6%

C9. What is your annual income from this child care setting? Is it . . .

MARK ONE ONLY

29.1% 1 ☐ Less than $15,000

35.6% 2 ☐ $15,000 to $24,999

26.2% 3 ☐ $25,000 to $49,999

1.9% 4 ☐ $50,000 to $74,999

1.1% 5 ☐ $75,000 to 150,000

0.1% 6 ☐ $150,000 or more

Missing 5.8%

Not applicable 0.2%

C10. As part of your employment does your child care setting offer any of the following benefits?

MARK ONE PER ROW

	YES	NO	DON'T KNOW	MISSING	N
a. Retirement/pension plan.....	1 ☐ 40.0%	0 ☐ 44.6%	d ☐ 10.1%	**5.2%**	**0.2%**
b. Life insurance.....	1 ☐ 41.0%	0 ☐ 44.2%	d ☐ 9.7%	**4.8%**	**0.2%**
c. Paid family/maternity leave.....	1 ☐ 30.4%	0 ☐ 48.6%	d ☐ 15.4%	**5.4%**	**0.3%**
d. Paid health insurance	1 ☐ 41.8%	0 ☐ 46.5%	d ☐ 6.4%	**5.1%**	**0.2%**
e. Dental insurance.....	1 ☐ 48.3%	0 ☐ 41.7%	d ☐ 5.0%	**4.8%**	**0.2%**
f. Paid sick leave.....	1 ☐ 60.2%	0 ☐ 31.3%	d ☐ 4.1%	**4.1%**	**0.2%**
g. Paid holidays.....	1 ☐ 76.8%	0 ☐ 17.0%	d ☐ 2.8%	**3.3%**	**0.2%**
h. Paid vacations.....	1 ☐ 67.3%	0 ☐ 25.0%	d ☐ 4.1%	**3.4%**	**0.2%**
i. Free or reduced child care for your own child(ren)	1 ☐ 43.0%	0 ☐ 30.6%	d ☐ 18.4%	**7.6%**	**0.4%**
j. Anything else? (*please specify*) _____	1 ☐ 4.5%	0 ☐ 24.3%	d ☐ 13.8%	**57.2%**	**0.2%**

D. EDUCATION AND EXPERIENCE

D0. Do you currently hold a Child Development Associate (CDA) credential?

34.9% 1 ☐ Yes

61.8% 0 ☐ No

Missing 3.3%

D1. What is the highest level of education you have completed?
MARK ONE ONLY

17.8% 0 ☐ High school diploma or GED

32.3% 1 ☐ College course(s) without a degree

19.3% 2 ☐ Associate's degree

21.8% 3 ☐ Bachelor's degree

2.7% 4 ☐ Master's degree

0.9% 5 ☐ Education specialist or professional diploma based on at least one year of course work past a Master's degree level

0.1% 6 ☐ Doctorate

2.5% 7 ☐ Other (*please specify)* _____

Additional code based on frequent "Other" specified

1.4% 8 ☐ Certificate

Missing 1.1%

D2. In what field did you obtain your highest degree?
MARK ALL THAT APPLY

11.8% 1 ☐ Child development or developmental psychology

40.3% 2 ☐ Early childhood education

5.8% 3 ☐ Elementary education

2.3% 4 ☐ Special education

31.0% 5 ☐ Other (*please specify)* _____

Missing 18.1%

D3. How many college courses have you completed in the following areas?

MARK ONE PER ROW

	0 None	1	2	3	4	5	6 or more	MISSING
a. Early childhood education	0 ☐ 26.3%	1 ☐ 9.6%	2 ☐ 5.3%	3 ☐ 5.9%	4 ☐ 3.9%	5 ☐ 1.7%	6 ☐ 35.6%	**11.8%**
b. Elementary education	0 ☐ 53.3%	1 ☐ 5.7%	2 ☐ 3.2%	3 ☐ 2.9%	4 ☐ 1.3%	5 ☐ 0.5%	6 ☐ 9.5%	**23.7%**
c. Special education	0 ☐ 47.8%	1 ☐ 13.0%	2 ☐ 6.6%	3 ☐ 2.8%	4 ☐ 1.5%	5 ☐ 0.9%	6 ☐ 3.9%	**23.5%**
d. English as a second language (ESL)	0 ☐ 60.2%	1 ☐ 5.0%	2 ☐ 2.2%	3 ☐ 1.4%	4 ☐ 0.8%	5 ☐ 0.6%	6 ☐ 4.8%	**24.9%**
e. Child development	0 ☐ 24.3%	1 ☐ 13.6%	2 ☐ 8.9%	3 ☐ 5.8%	4 ☐ 4.1%	5 ☐ 1.6%	6 ☐ 27.1%	**14.7%**
f. Infant development	0 ☐ 31.6%	1 ☐ 16.5%	2 ☐ 10.4%	3 ☐ 6.5%	4 ☐ 2.8%	5 ☐ 1.2%	6 ☐ 12.7%	**18.0%**
g. Methods of teaching reading	0 ☐ 47.4%	1 ☐ 12.8%	2 ☐ 7.4%	3 ☐ 3.8%	4 ☐ 2.1%	5 ☐ 0.8%	6 ☐ 4.5%	**21.2%**
h. Methods of teaching mathematics	0 ☐ 50.8%	1 ☐ 13.5%	2 ☐ 6.1%	3 ☐ 2.7%	4 ☐ 1.8%	5 ☐ 0.7%	6 ☐ 3.1%	**21.4%**
i. Methods of teaching science	0 ☐ 52.2%	1 ☐ 13.8%	2 ☐ 5.4%	3 ☐ 2.3%	4 ☐ 1.8%	5 ☐ 0.3%	6 ☐ 3.1%	**21.3%**

D4. Including this year, how many years have you worked with infants and/or toddlers?

|__|__| YEARS

N = 746

MEDIAN 6.0; MEAN 8.6

RANGE 0-40

MISSING 23.3%

E. DEMOGRAPHIC INFORMATION

E1. Are you…

97.0% 1 ☐ Female

2.5% 2 ☐ Male

Missing 0.5%

E2. In what year were you born?

|__|__|__|__| YEAR

N = 952

MEDIAN 1976

RANGE 1935-1995

E3. What is your first language?

MARK ONE ONLY

77.1% 1 ☐ English

15.8% 2 ☐ Spanish

6.5% 3 ☐ Other (*please specify*) _____

Missing 0.6%

E4. Please indicate any other languages you speak fluently.

SELECT ONE OR MORE

29.4% 0 ☐ No other language spoken

19.1% 1 ☐ English 19.1%

7.2% 2 ☐ Spanish

5.5% 3 ☐ Other (*please specify*) _____

Additional code based on frequent "Other" specified

1.1% 3 ☐ Sign Language

Missing 3.5%

E5. Are you of Spanish, Hispanic or Latino origin?

27.0% 1 ☐ Yes

71.9% 0 ☐ No

Missing 1.1%

E6. What is your race?

SELECT ONE OR MORE

53.1% 1 ☐ White

30.9% 2 ☐ Black or African-American

3.6% 3 ☐ Asian

3.0% 4 ☐ American Indian or Alaskan Native

0.4% 5 ☐ Native Hawaiian or other Pacific Islander

Missing 11.8%

Thank you for your participation. If you have any questions about this questionnaire or the Q-CCIIT project, please call the survey director, Shannon Monahan, at (609) 275-2207.

Please return this questionnaire in the envelope provided. If you no longer have the envelope, please mail this questionnaire to:

Mathematica Policy Research

Attn: Receipt Control – Project 06861

P.O. Box 2393

Princeton, NJ 08543-2393

This page has been left blank for double-sided copying.

APPENDIX B

SUBGROUP DESCRIPTIVE STATISTICS FOR THE ORCE, ITERS-R, AND FCCERS-R

This page has been left blank for double-sided copying.

Table B.1. Descriptive Statistics for the ORCE Scales: By Program Type

	EHS				Other Center				FCC			
Scales	N	M (SD)	Min.	Max.	N	M (SD)	Min.	Max.	N	M (SD)	Min.	Max.
Overall Qualitative Rating	20	3.22 (0.34)	2.57	3.80	58	3.19 (0.43)	1.90	3.85	41	3.24 (0.48)	1.79	3.95
Sensitivity/Responsiveness to Distress[a]	14	3.51 (0.82)	1.00	4.00	41	3.58 (0.57)	2.00	4.00	26	3.53 (0.68)	1.40	4.00
Sensitivity/Responsiveness to Non-Distress	20	3.25 (0.54)	2.50	4.00	58	3.31 (0.55)	1.10	4.00	41	3.34 (0.64)	1.50	4.00
Lack of Intrusiveness	20	3.90 (0.19)	3.19	4.00	58	3.87 (0.26)	2.75	4.00	41	3.82 (0.42)	2.13	4.00
Lack of Detachment/Disengagement	20	3.72 (0.44)	2.46	4.00	58	3.66 (0.54)	1.92	4.00	41	3.71 (0.45)	2.38	4.00
Stimulation of Cognitive Development	20	2.23 (0.70)	1.19	3.44	58	1.97 (0.58)	1.17	3.38	41	2.28 (0.65)	1.00	3.75
Positive Regard for the Child	20	3.44 (0.40)	2.65	4.00	58	3.48 (0.46)	2.33	4.00	41	3.44 (0.51)	2.25	4.00
Lack of Negative Regard for the Child	20	4.00 (0.01)	3.96	4.00	58	3.99 (0.05)	3.63	4.00	41	3.95 (0.15)	3.38	4.00
Lack of Flatness of Affect	20	3.79 (0.40)	2.79	4.00	58	3.71 (0.52)	1.75	4.00	41	3.74 (0.54)	2.00	4.00
Fostering Exploration[b]	8	2.31 (0.50)	1.50	3.13	12	2.32 (0.89)	1.00	3.75	20	2.46 (1.01)	1.00	4.00
Positive Rating	20	2.89 (0.45)	2.32	3.67	58	2.88 (0.48)	1.54	3.75	41	2.96 (0.58)	1.47	3.92
Lack of Negative Rating	20	3.85 (0.22)	3.25	4.00	58	3.81 (0.29)	2.83	4.00	41	3.80 (0.31)	2.78	4.00
Language Stimulation	20	2.17 (0.75)	1.03	3.75	58	1.74 (0.64)	0.66	3.55	41	2.17 (0.79)	0.63	4.13
Positive Behavior Toward Child	20	0.72 (0.56)	0.17	2.08	58	0.62 (0.55)	0.04	2.70	41	0.56 (0.51)	0.00	2.44
Negative Behavior Toward Child	20	0.20 (0.33)	0.00	1.21	58	0.32 (0.45)	0.00	1.53	41	0.27 (0.38)	0.00	1.39

Source: Q-CCIIT Fall 2012 Psychometric Field Test.

[a] For 6- and 15-month form.

[b] For 36-month form only.

Table B.2. Descriptive Statistics for the ITERS-R Subscales, by Concentration of DLLS

	High				Low			
Subscale	Mean	SD	Min	Max	Mean	SD	Min	Max
Personal Care	1.61**	0.46	1	2.5	2.18	0.65	1	3.67
Space and Furnishings	3.58	0.80	1.6	4.6	3.80	1.04	2.2	6
Listening and Talking	3.78	1.66	1.33	7	3.54	1.28	1.33	6
Activities	3.11	1.02	1.5	5.22	3.03	0.92	1.25	5.2
Interaction/Social	4.81	1.57	2.25	7	4.12	1.49	1.25	7
Program Structure	3.65	1.40	1	5.5	3.46	1.38	1	7
Total	3.27	0.79	1.62	4.7	3.24	0.83	1.78	5.07
Child/adult Ratio	3.23	1.07	1.5	5.33	4.19	1.95	1.8	14
Sample Size	**17**				**45-46**			

Source: Q-CCIIT Fall 2012 Psychometric Field Test Data.

Note: Item ratings range from 1 to 7.

Table B.3. Descriptive Statistics for the FCCER-R Subscales, by Concentration of DLLS

	High				Low			
Subscales	Mean	SD	Min	Max	Mean	SD	Min	Max
Personal Care	1.90	0.70	1.17	3.83	1.90	0.63	1	3.67
Space and Furnishings	2.77	1.04	1.33	4.67	2.71	0.82	1.17	4.50
Listening and Talking	3.24	0.91	1.33	4.67	3.75	1.40	1	6.67
Activities	2.46	0.61	1.55	3.4	2.42	0.85	1.09	3.82
Interaction/Social	4.39	1.49	1	5.75	4.22	1.79	1	7
Program Structure	3.34	1.20	1.67	5.33	3.31	1.59	1	6.67
Total	2.81	0.73	1.53	3.91	2.81	0.82	1.12	3.91
Child/adult Ratio	4.36	1.92	2	8	5.05	2.62	2	14
Sample Size	**14**				**32-33**			

Source: Q-CCIIT Fall 2012 Psychometric Field Test Data.

Note: Item ratings range from 1 to 7.

This page has been left blank for double-sided copying.

APPENDIX C

Q-CCIIT ITEM DESCRIPTIVE STATISTICS

This page has been left blank for double-sided copying.

Table C.1. Types of Interactions, by Classroom Type

	Full Sample		Infant		Toddler		FCC		Center	
Type of Interactions	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Free play	0.38	0.26	0.52	0.26	0.28	0.20	0.35	0.25	0.39	0.26
Book reading	0.09	0.12	0.08	0.12	0.09	0.11	0.11	0.12	0.08	0.12
Book sharing (Goodnight Gorilla)	0.17	0.05	0.17	0.06	0.17	0.04	0.16	0.03	0.17	0.05
Circle time	0.1	0.14	0.04	0.09	0.13	0.14	0.12	0.16	0.09	0.13
Feeding/meal time	0.24	0.18	0.35	0.19	0.18	0.13	0.19	0.15	0.26	0.18
Outside free play	0.1	0.15	0.04	0.09	0.17	0.17	0.1	0.14	0.11	0.15
Diapering/toileting	0.1	0.13	0.16	0.16	0.06	0.1	0.09	0.12	0.11	0.14
Transition	0.18	0.19	0.15	0.19	0.20	0.18	0.18	0.19	0.18	0.19
Small group teacher-directed activity	0.15	0.18	0.09	0.15	0.16	0.16	0.21	0.21	0.13	0.16
Other	0.06	0.12	0.09	0.16	0.03	0.09	0.05	0.10	0.06	0.13
Sample Size	**400**		**136**		**154**		**110**		**290**	

Source: Q-CCIIT Fall 2012 Psychometric Field Test Data.

Note: The item score is the mean across valid observation cycles (5 minutes or longer), which represents the percentage of observation cycles during which the type of interaction is observed.

Table C.2. Types of Interactions, by Concentration of DLLs

	High		Low	
Type of Interactions	Mean	SD	Mean	SD
Free play	0.37	0.25	0.39	0.27
Book reading	0.08	0.11	0.10	0.13
Book sharing (Goodnight Gorilla)	0.16	0.05	0.17	0.05
Circle time	0.10	0.13	0.10	0.14
Feeding/meal time	0.22	0.15	0.25	0.19
Outside free play	0.12	0.16	0.10	0.14
Diapering/toileting	0.09	0.13	0.10	0.14
Transition	0.15	0.17	0.19	0.19
Small group teacher-directed activity	0.16	0.18	0.14	0.18
Other	0.05	0.12	0.06	0.12
Sample Size	**126**		**253**	

Source: Q-CCIIT Fall 2012 Psychometric Field Test Data.

Note: The item score is the mean across valid observation cycles (5 minutes or longer), which represents the percentage of observation cycles during which the type of interaction is observed.

Table C.3. Types of Talk, by Classroom Type

	Full Sample		Infant		Toddler		FCC		Center	
Type of Talk	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Parallel language	0.82	0.53	0.86	0.53	0.82	0.53	0.76	0.53	0.84	0.53
Descriptive talk	0.73	0.47	0.64	0.49	0.78	0.45	0.77	0.45	0.71	0.47
Songs/poems	0.42	0.34	0.45	0.37	0.43	0.3	0.36	0.34	0.44	0.33
Explanations	0.29	0.32	0.21	0.29	0.32	0.33	0.34	0.34	0.27	0.32
Reasoning	0.05	0.15	0.04	0.12	0.05	0.14	0.07	0.18	0.04	0.13
Anticipatory talk	0.48	0.36	0.46	0.38	0.51	0.35	0.46	0.33	0.49	0.37
Decontextualized	0.33	0.31	0.26	0.28	0.37	0.33	0.36	0.32	0.32	0.31
Sample Size	**400**		**136**		**154**		**110**		**290**	

Source: Q-CCIIT Fall 2012 Psychometric Field Test Data.

Note: The item score is the mean across valid observation cycles (5 minutes or longer). The item ratings are 0 (none), 1 (once), and 2 (more than once).

Table C.4. Types of Talk, by Concentration of DLLs

	High		Low	
Type of Talk	Mean	SD	Mean	SD
Parallel language	0.86	0.59	0.80	0.50
Descriptive talk	0.68	0.46	0.77	0.47
Songs/poems	0.49	0.37	0.39	0.32
Explanations	0.32	0.33	0.27	0.32
Reasoning	0.06	0.18	0.05	0.13
Anticipatory talk	0.48	0.34	0.48	0.37
Decontextualized	0.31	0.33	0.34	0.31
Sample Size	**126**		**253**	

Source: Q-CCIIT Fall 2012 Psychometric Field Test Data.

Note: The item score is the mean across valid observation cycles (5 minutes or longer). The item ratings are 0 (none), 1 (once), and 2 (more than once).

Table C.5. Concept Development, by Classroom Type

	Full Sample		Infant		Toddler		FCC		Center	
Concept	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Color	0.26	0.22	0.17	0.18	0.3	0.22	0.32	0.24	0.24	0.22
Shape	0.06	0.12	0.02	0.05	0.06	0.11	0.12	0.15	0.04	0.09
Letters	0.09	0.16	0.05	0.11	0.08	0.12	0.16	0.21	0.07	0.12
Numbers	0.24	0.24	0.17	0.21	0.28	0.24	0.28	0.25	0.23	0.23
Size	0.07	0.12	0.05	0.11	0.06	0.11	0.09	0.13	0.06	0.11
More/less	0.03	0.08	0.03	0.09	0.02	0.07	0.02	0.08	0.03	0.08
Same/different	0.01	0.03	0	0.02	0.01	0.04	0.01	0.03	0.00	0.03
Sounds	0.02	0.07	0.02	0.08	0.01	0.05	0.03	0.09	0.02	0.07
Other sensory	0.11	0.18	0.12	0.18	0.1	0.17	0.09	0.19	0.11	0.17
Categories (animals, furniture, plants, etc.)	0.18	0.17	0.15	0.14	0.2	0.19	0.2	0.17	0.17	0.17
Feelings/States	0.22	0.22	0.30	0.25	0.18	0.18	0.18	0.18	0.24	0.22
On/off	0.01	0.05	0.01	0.04	0.02	0.06	0.01	0.03	0.01	0.05
Up/down	0.07	0.13	0.06	0.12	0.09	0.14	0.05	0.1	0.08	0.13
In/out	0.04	0.09	0.03	0.08	0.03	0.07	0.05	0.11	0.03	0.07
Other spatial concepts	0.03	0.08	0.01	0.06	0.05	0.09	0.04	0.09	0.03	0.08
Other opposites	0.04	0.10	0.03	0.09	0.05	0.09	0.05	0.11	0.04	0.09
Body parts	0.00	0.03	0.00	0.02	0.00	0.03	0.00	0.03	0.00	0.03
Literacy concepts	0.00	0.03	0.00	0.00	0.00	0.03	0.01	0.04	0.00	0.03
Other	0.01	0.03	0.01	0.04	0.00	0.03	0.01	0.04	0.00	0.03
Number of unique concepts presented across cycles	5.28	2.59	4.29	2.48	5.75	2.51	5.83	2.5	5.07	2.60
Sample Size	**400**		**136**		**154**		**110**		**290**	

Source: Q-CCIIT Fall 2012 Psychometric Field Test Data.

Note: The item score is the mean across valid observation cycles (5 minutes or longer), which represents the percentage of observation cycles during which the concept is presented or discussed.

Table C.6. Concept Development, by Concentration of DLLS

	High		Low	
Concept	Mean	SD	Mean	SD
Color	0.31	0.24	0.24	0.22
Shape	0.08	0.14	0.06	0.10
Letters	0.11	0.18	0.09	0.15
Numbers	0.29	0.26	0.22	0.23
Size	0.07	0.12	0.06	0.11
More/less	0.03	0.09	0.02	0.07
Same/different	0.01	0.04	0.00	0.03
Sounds	0.01	0.05	0.03	0.08
Other sensory	0.11	0.20	0.10	0.17
Categories (animals, furniture, plants, etc.)	0.16	0.16	0.19	0.18
Feelings/States	0.19	0.22	0.23	0.22
On/off	0.01	0.05	0.01	0.05
Up/down	0.09	0.15	0.06	0.11
In/out	0.03	0.08	0.03	0.08
Other spatial concepts	0.04	0.09	0.03	0.08
Other opposites	0.06	0.11	0.04	0.09
Body parts	0.00	0.03	0.00	0.03
Literacy concepts	0.00	0.03	0.00	0.03
Other	0.00	0.03	0.01	0.03
Number of unique concepts presented across cycles	5.44	2.83	5.21	2.45
Sample Size	**126**		**253**	

Source: Q-CCIIT Fall 2012 Psychometric Field Test Data.

Note: The item score is the mean across valid observation cycles (5 minutes or longer), which represents the percentage of observation cycles during which the concept is presented or discussed.

Table C.7. Support for Social-Emotional Development, by Classroom Type

	Infant				Toddler				FCC				Center			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Responding contingently to distress[a]	4.67^^	1.29	1.00	7.00	4.64**	1.43	1.00	7.00	4**	1.54	1.00	7.00	4.66**	1.35	1	7
Responding to social cues[a]	4.57^^*	1.11	1.83	7.00	4.88**	1	1.83	6.67	4.44	1.19	1.17	7.00	4.73*	1.06	1.83	7
Responding to emotional cues[a]	4.56	1.12	1.83	7.00	4.71*	1.03	1.33	6.83	4.38	1.23	1.17	7.00	4.64*	1.07	1.33	7
Builds a positive relationship[a]	5.03	1.11	2.17	7.00	4.91	0.99	1.17	6.67	4.72*	1.19	1.00	6.83	4.97*	1.05	1.17	7
Supporting peer interaction/play[a]	3.08^^**	1.13	1.00	6.33	3.48	1.2	1.00	6.83	3.44*	1.2	1.00	6.50	3.29	1.18	1	6.83
Support for social problem solving[b]	3.15^^^***	1.21	1.00	6.00	3.93	1.16	1.00	6.00	3.65*	1.42	1.00	7.00	3.65	1.24	1	6
Responsive routines[b]	4.11^^**	1.76	1.00	7.00	4.66*	1.4	1.00	7.00	4.25	1.63	1.00	7.00	4.40	1.60	1	7
Classroom limits and management[b]	3.84^*	1.58	1.00	7.00	4.2*	1.27	2.00	7.00	3.78	1.4	1.00	7.00	4.05	1.42	1	7
Sense of belonging[b]	4.3^^^***	1.5	1.00	7.00	4.92***	1.36	1.00	7.00	3.96	1.48	1.00	7.00	4.63***	1.46	1	7
Sample Size[c]	**74-136**				**94-154**				**61-110**				**205-290**			

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Item ratings range from 1 to 7.

The results of the test for infant vs toddler are on the infant mean
The results for toddler vs FCC are on the toddler mean
The results for FCC vs infant are on the FCC mean

[a]The item score is the mean across valid observation cycles (five minutes or longer).

[b]Rated across the visit.

[c]Two items, responding contingently to distress and classroom limits and management have smaller sample sizes due to use of “not applicable”. If no distress occurred, the item was noted as ‘not applicable” and the sample size for that item was reduced. Additionally, classroom limits and management was only coded for mobile children so the sample size for that item is lower.

Table C.8. Support for Social-Emotional Development, Full Sample and by Concentration of DLLs

	Full Sample				High				Low			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Responding contingently to distress[a]	4.51	1.42	1	7	4.35	1.40	1	7	4.52	1.41	1	7
Responding to social cues[a]	4.65	1.1	1.17	7	4.72	1.09	1.33	7	4.63	1.10	1.17	7
Responding to emotional cues[a]	4.57	1.12	1.17	7	4.63	1.08	1.17	7	4.56	1.13	1.83	7
Builds a positive relationship[a]	4.9	1.09	1	7	4.91	1.06	1	7	4.92	1.08	1.67	7
Supporting peer interaction/play[a]	3.33	1.19	1	6.83	3.46	1.15	1	6.5	3.30	1.21	1	6.83
Support for social problem solving[b]	3.65	1.29	1	7	3.85	1.34	1	7	3.58	1.27	1	7
Responsive routines[b]	4.36	1.61	1	7	4.49	1.67	1	7	4.31	1.60	1	7
Classroom limits and management[b]	3.97	1.42	1	7	4.03	1.50	1	7	3.95	1.35	1	7
Sense of belonging[b]	4.45	1.49	1	7	4.47	1.63	1	7	4.45	1.40	1	7
Sample Size[c]	**266-400**				**77-126**				**176-253**			

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Item ratings range from 1 to 7.

[a]The item score is the mean across valid observation cycles (five minutes or longer).

[b]Rated across the visit.

[c]Two items, responding contingently to distress and classroom limits and management have smaller sample sizes due to use of “not applicable”. If no distress occurred, the items was noted as ‘not applicable” and the sample size for that item was reduced. Additionally, classroom limits and management was only coded for mobile children so the sample size for that item is lower.

Table C.9. Support for Cognitive Development, by Classroom Type

	Infant				Toddler				FCC				Center			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Supporting object exploration[a]	3.64^^**	1.1	1.5	6.67	4.1*	1.24	1	7	3.74	1.34	1	7	3.88	1.20	1	7
Scaffolding problem solving[a]	2.72^^^***	1.38	1	6	3.34	1.37	1	6	3.08*	1.43	1	6.67	3.06	1.41	1	6
Giving choices[b]	3.06^^^***	1.57	1	7	3.97*	1.41	1	7	3.6**	1.54	1	7	3.57	1.55	1	7
Extending pretend play[b]	2.61^^^***	1.48	1	6	3.38	1.67	1	7	3.34***	1.68	1	7	3.05	1.64	1	7
Explicit teaching[b]	2.86^^^***	1.33	1	7	3.56	1.22	1	7	3.52***	1.58	1	7	3.23	1.32	1	7
Supervises or joins in play and activities[b]	4.37^^***	1.7	1	7	5.05*	1.36	2	7	4.62	1.7	1	7	4.73	1.56	1	7
Sample Size	**112-136**				**152-154**				**108-110**				**264-290**			

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Item ratings range from 1 to 7.

The results of the test for infant vs toddler are on the infant mean

The results for toddler vs FCC are on the toddler mean

The results for FCC vs infant are on the FCC mean

[a]The item score is the mean across valid observation cycles (five minutes or longer).

[b]Rated across the visit.

Table C.10. Support for Cognitive Development, Full Sample and by Concentration of DLLS

	Full Sample				High				Low			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Supporting object exploration[a]	3.84	1.24	1	7	4.02	1.34	1	7	3.79	1.20	1	7
Scaffolding problem solving[a]	3.06	1.41	1	6.67	3.24	1.49	1	6.5	2.99	1.39	1	6.67
Giving choices[b]	3.58	1.54	1	7	3.74	1.60	1	6	3.53	1.54	1	7
Extending pretend play[b]	3.14	1.65	1	7	3.44*	1.67	1	7	3.00	1.64	1	7
Explicit teaching[b]	3.31	1.4	1	7	3.37	1.47	1	7	3.26	1.39	1	7
Supervises or joins in play and activities[b]	4.7	1.6	1	7	4.82	1.68	1	7	4.67	1.57	1	7
Sample Size	**372-399**				**118-126**				**234-253**			

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Item ratings range from 1 to 7.

[a]The item score is the mean across valid observation cycles (five minutes or longer).

[b]Rated across the visit.

Table C.11. Support for Language Development, by Classroom Type

	Infant				Toddler				FCC				Center			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Caregiver use of varied vocabulary[a]	3.7^^**	1.16	1	6.5	4.11*	1.02	1.2	6.17	3.83	1.29	1	6.67	3.92	1.11	1	6.5
Conversational turn-taking[a]	3.32^^^***	1.37	1	6.67	4.05	1.17	1.17	6.4	3.84**	1.37	1	7	3.71	1.32	1	6.67
Use of questions[a]	3.09^^^***	0.94	1.5	6	3.71	0.93	1.5	6.33	3.56***	1.09	1.33	7	3.42	0.98	1.5	6.33
Extending children's language use[a]	3^^^***	1.2	1	5.83	3.58	1.1	1	6	3.49**	1.22	1	7	3.30	1.19	1	6
Engaging children in books[a]	4.62^^^***	1.39	1	7	5.25	1.44	1	7	5.05*	1.34	2	7	4.95	1.45	1	7
Variety of words[a]	4.15^^^***	1.2	2	7	4.8	1.28	1	7	4.67**	1.3	1	7	4.50	1.29	1	7
Variety of types of sentences[a]	3.6^^^***	1.24	1	7	4.19	1.07	1	7	4.13***	1.19	1	7	3.92	1.19	1	7
Features of talk[b]	4.6^^^***	1.51	1	7	5.31**	1.24	1	7	4.85	1.53	1	7	4.98	1.42	1	7
Talk about things not present[b]	2.88^^^***	1.4	1	7	3.53	1.38	1	7	3.33*	1.66	1	7	3.23	1.43	1	7
Positive attitude toward books[b]	4.07^^^***	1.48	1	7	4.86	1.44	2	7	4.58**	1.59	1	7	4.49	1.51	1	7
Sample Size	**132-136**				**150-154**				**107-110**				**282-290**			

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Item ratings range from 1 to 7.

The results of the test for infant vs toddler are on the infant mean
The results for toddler vs FCC are on the toddler mean
The results for FCC vs infant are on the FCC mean

[a]The item score is the mean across valid observation cycles (five minutes or longer).

[b]Rated across the visit.

Table C.12. Support for Language Development, Full Sample and by Concentration of DLLS

	Full Sample				High				Low			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Caregiver use of varied vocabulary[a]	3.89	1.16	1	6.67	3.96	1.23	1	6.67	3.89	1.14	1	6.67
Conversational turn-taking[a]	3.74	1.33	1	7	3.89	1.46	1.17	7	3.68	1.26	1	7
Use of questions[a]	3.46	1.01	1.33	7	3.62*	1.07	1.33	6	3.40	0.98	1.5	7
Extending children's language use[a]	3.35	1.2	1	7	3.45	1.28	1	7	3.31	1.16	1	7
Engaging children in books[a]	4.98	1.42	1	7	5.14	1.41	1	7	4.90	1.42	1	7
Variety of words[a]	4.55	1.29	1	7	4.63	1.37	1	7	4.51	1.26	1	7
Variety of types of sentences[a]	3.98	1.19	1	7	4.04	1.27	1	7	3.93	1.15	1	7
Features of talk[b]	4.94	1.45	1	7	4.86	1.57	1	7	5.00	1.36	1	7
Talk about things not present[b]	3.26	1.49	1	7	3.31	1.54	1	7	3.23	1.45	1	7
Positive attitude toward books[b]	4.52	1.53	1	7	4.46	1.51	2	7	4.56	1.55	1	7
Sample Size	**389-400**				**121-126**				**247-253**			

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Item ratings range from 1 to 7.

[a]The item score is the mean across valid observation cycles (five minutes or longer).

[b]Rated across the visit.

Table C.13. The Average Percentage of Classrooms Offering Activities by Classroom Type

	Infant	Toddler	FCC	Center
Gross motor activities (e.g., playing outside, walking, crawling)	76.47	84.42	66.36	80.69
Fine motor activities (e.g., infants—nested objects, ring toys; toddlers—puzzles, beads, blocks to put together, shake)	94.85	92.86	92.73	93.79
Art and crafts (e.g., painting)	19.12	61.04	53.64	41.38
Music (e.g., singing, playing instruments, dancing)	80.88	93.51	80.91	87.59
Sensory play (e.g., water, sand)	26.47	56.49	30.91	42.41
Dramatic play	38.24	77.27	76.36	58.97
Books/storytelling (do not include *Goodnight Gorilla*)	70.59	81.82	78.18	76.55
Other	0.74	1.95	1.82	1.38
Sample Size	**136**	**154**	**110**	**290**

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: The table shows the percentages of classrooms that offered the activities.

Table C.14. The Average Percentage of Classrooms Offering Activities Full Sample and by Concentration of DLLS

	Full Sample	High	Low
Gross motor activities (e.g., playing outside, walking, crawling)	76.75	76.98	76.68
Fine motor activities (e.g., infants—nested objects, ring toys; toddlers—puzzles, beads, blocks to put together, shake)	93.5	93.65	94.07
Art and crafts (e.g., painting)	44.75	47.62	42.69
Music (e.g., singing, playing instruments, dancing)	85.75	84.13	86.17
Sensory play (e.g., water, sand)	39.25	45.24	36.36
Dramatic play	63.75	66.67	62.06
Books/storytelling (do not include *Goodnight Gorilla*)	77	75.40	77.47
Other	1.5	1.59	1.58
Sample Size	**400**	**126**	**253**

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: The table shows the percentages of classrooms that offered the activities.

Table C.15. Areas of Concern: Frequency, by Classroom Type

	Infant		Toddler		FCC		Center	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Physically harsh	1.06	0.24	1.07	0.25	1.11	0.31	1.06	0.24
Verbally harsh	1.02	0.15	1.07	0.28	1.18	0.43	1.05	0.23
Restricts children (other than safety)	1.13	0.44	1.09	0.3	1.25	0.58	1.11	0.37
Communication mismatch (between tone/facial expression and message)	1.05	0.22	1.03	0.16	1.13	0.39	1.04	0.19
Repeatedly singles out a child	1.05	0.22	1.1	0.35	1.16	0.46	1.08	0.30
Ignores children	1.44	0.64	1.28	0.52	1.44	0.61	1.36	0.59
Children unoccupied	1.51	0.64	1.39	0.6	1.54	0.67	1.44	0.62
Overwhelms children	1.01	0.12	1.03	0.18	1.1	0.36	1.02	0.15
Children stressed by demands	1.01	0.12	1.06	0.24	1.07	0.29	1.04	0.19
Adult television	1	0	1	0	1.1	0.41	1.00	0.00
Children watch children's media	1.01	0.17	1.02	0.14	1.25	0.53	1.02	0.16
Supervision of safety is poor	1.1	0.34	1.15	0.39	1.37	0.62	1.13	0.37
Unsafe environment	1.07	0.25	1.11	0.32	1.33	0.64	1.09	0.29
General health provisions not available/sanitary practices not followed	1.1	0.31	1.13	0.38	1.4	0.65	1.12	0.34
Level of chaos	2.38	1.13	2.71	1.41	2.87	1.41	2.55	1.29
Sample Size	**134-135**		**152-153**		**107-110**		**286-288**	

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Item ratings range from 1 to 3 except for "level of chaos," for which the ratings range from 1 to 7.

Table C.16. Areas of Concern: Frequency, Full Sample and by Concentration of DLLs

	Full Sample		High		Low	
	Mean	SD	Mean	SD	Mean	SD
Physically harsh	1.08	0.26	1.09	0.28	1.07	0.26
Verbally harsh	1.09	0.31	1.05	0.21	1.10	0.34
Restricts children (other than safety)	1.15	0.44	1.08	0.33	1.19	0.50
Communication mismatch (between tone/facial expression and message)	1.06	0.26	1.06	0.25	1.06	0.28
Repeatedly singles out a child	1.1	0.35	1.03	0.18	1.13	0.40
Ignores children	1.38	0.59	1.38	0.58	1.38	0.61
Children unoccupied	1.47	0.64	1.43	0.56	1.48	0.67
Overwhelms children	1.05	0.23	1.05	0.25	1.03	0.20
Children stressed by demands	1.05	0.22	1.04	0.20	1.05	0.23
Adult television	1.03	0.22	1.04	0.27	1.02	0.20
Children watch children's media	1.08	0.32	1.06	0.29	1.10	0.35
Supervision of safety is poor	1.19	0.47	1.12	0.35	1.22	0.50
Unsafe environment	1.16	0.43	1.17	0.43	1.14	0.42
General health provisions not available/sanitary practices not followed	1.2	0.47	1.17	0.42	1.21	0.49
Level of chaos	2.64	1.33	2.52	1.32	2.63	1.27
Sample Size	**394-398**		**124-126**		**249-251**	

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Item ratings range from 1 to 3 except for "level of chaos," for which the ratings range from 1 to 7.

Table C.17. Percentage of Classrooms with Extreme Areas of Concern, by Classroom Type

	Infant	Toddler	FCC	Center
Physically harsh	0.00	0.00	3.64	0.00
Verbally harsh	0.00	0.00	0.00	0.00
Restricts children (other than safety)	1.48	0.00	2.73	0.69
Communication mismatch (between tone/facial expression and message)	0.00	0.00	0.00	0.00
Ignores children	0.74	0.65	3.64	0.69
Adult television	0.00	0.00	0.91	0.00
Children watch children's media	.74	1.32	8.18	1.05
Supervision of safety is poor	0.00	0.00	4.55	0.00
Unsafe environment	0.00	0.00	0.00	0.00
General health provisions not available/sanitary practices not followed	0.00	0.00	2.73	0.00
Sample Size	**135**	**152-153**	**109-110**	**287-288**

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Table shows the percentages of classrooms that had extreme concerns.

Table C.18. Percentage of Classrooms with Extreme Areas of Concern, Full Sample and by Concentration of DLLs

	Full Sample	High	Low
Physically harsh	1.01	0.00	1.59
Verbally harsh	0.00	0.00	0.00
Restricts children (other than safety)	1.26	1.59	1.20
Communication mismatch (between tone/facial expression and message)	0.00	0.00	0.00
Ignores children	1.51	1.59	1.20
Adult television	0.25	0.00	0.40
Children watch children's media	3.02	1.59	4.00
Supervision of safety is poor	1.26	0.79	1.59
Unsafe environment	0.00	0.00	0.00
General health provisions not available/sanitary practices not followed	0.75	1.59	0.40
Sample Size	**397-398**	**125-126**	**250-251**

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Table shows the percentages of classrooms that had extreme concerns.

Table C.19. Environmental Items, by Classroom Type

	Infant		Toddler		FCC		Center	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Schedule balances types of activities	3.34	0.88	3.54	0.73	3.19	0.92	3.44	0.80
Quiet area is available	2.87	1.33	3	1.3	2.06	1.34	2.94	1.32
Caregiving space is organized	3.7	0.62	3.81	0.47	3.23	0.93	3.76	0.55
Caregivers are supportive of parents	3.83	0.52	3.67	0.69	3.4	0.95	3.75	0.62
Sample Size	**95-136**		**109-153**		**58-110**		**204-289**	

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Item ratings range from 1 (not at all) to 4 (strongly characteristic).

Table C.20. Environmental Items, Full Sample and by Concentration of DLLs

	Full Sample		High		Low	
	Mean	SD	Mean	SD	Mean	SD
Schedule balances types of activities	3.37	0.85	3.28	0.90	3.43	0.82
Quiet area is available	2.7	1.38	2.98	1.27	2.59	1.40
Caregiving space is organized	3.61	0.71	3.59	0.77	3.63	0.68
Caregivers are supportive of parents	3.67	0.72	3.70	0.65	3.67	0.73
Sample Size	**262-398**		**77-126**		**173-251**	

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Item ratings range from 1 (not at all) to 4 (strongly characteristic).

Table C.21. Overall Level of Quality, by Classroom Type

	Infant		Toddler		FCC		Center	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Caregiver 1	4.81	1.33	5.11	1.11	4.67	1.29	4.97	1.22
Caregiver 2	4.64	1.19	4.77	1.34	4.07	1.42	4.71	1.27
Caregiver 3	4.36	1.26	4.47	1.08	4.08	1.63	4.41	1.18
Overall classroom	4.71	1.15	4.88	1.09	4.44	1.23	4.80	1.12
Sample Size	**67-135**		**66-151**		**25-109**		**133-286**	

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Ratings range from 1 (lowest quality) to 7 (highest quality).

Table C.22. Overall Level of Quality, Full Sample and by Concentration of DLLs

	Full Sample		High		Low	
	Mean	SD	Mean	SD	Mean	SD
Caregiver 1	4.89	1.25	5.06	1.25	4.83	1.23
Caregiver 2	4.59	1.32	4.64	1.43	4.58	1.25
Caregiver 3	4.36	1.26	4.39	1.36	4.40	1.20
Overall classroom	4.70	1.16	4.88	1.10	4.64	1.15
Sample Size	**158-395**		**158-395**		**95-251**	

Source: Q-CCIIT Fall 2012 Psychometric Field Test data.

Note: Ratings range from 1 (lowest quality) to 7 (highest quality).

APPENDIX D

TABLES FOR TEMPORAL STABILITY ON Q-CCIIT ITEMS

This page has been left blank for double-sided copying.

Table D.1. Test-Retest Correlations: Support for Social-Emotional Development Items

	Total	Infant	Toddler	FCC
Responding contingently to distress[a]	0.35***	0.05	0.19	0.45**
Responding to social cues[a]	0.71***	0.78***	0.38	0.74***
Responding to emotional cues[a]	0.76***	0.84***	0.49*	0.83***
Builds a positive relationship[a]	0.71***	0.78***	0.45	0.74***
Supervises or joins in play and activities[b]	0.65***	0.77***	0.54**	0.58***
Responsive routines[b]	0.72***	0.86***	0.36	0.70***
Classroom limits and management[b]	0.57***	0.68***	0.61**	0.21
Sense of belonging[b]	0.78***	0.83***	0.54**	0.74***
Sample Size	**61-62**	**18**	**14**	**29-30**

Source: Q-CCIIT Fall 2012 Psychometric Field Test

[a] The item rating used in analysis is the mean across valid observation cycles (five minutes or longer) rounded to the nearest integer.

[b] Rated across the visit.

*p<.05; **p<.01; ***p<.001.

Table D.2. Test-Retest Correlations: Support for Language and Literacy Development Items

	Total	Infant	Toddler	FCC
Caregiver use of varied vocabulary[a]	0.66***	0.75***	0.41	0.73***
Conversational turn-taking[a]	0.72***	0.82***	0.75***	0.63***
Use of questions[a]	0.82***	0.92***	0.69***	0.79***
Extending children's language use[a]	0.78***	0.75***	0.80***	0.76***
Engaging children in books[a]	0.33***	0.49**	0.54**	-0.02
Variety of words[a]	0.29**	0.44*	0.42	-0.10
Variety of types of sentences[a]	0.23*	0.54**	0.15	0.08
Features of talk[b]	0.75***	0.78***	0.40	0.78***
Talk about things not present[b]	0.53***	0.67***	0.56**	0.41**
Positive attitude toward books[b]	0.49***	0.70***	0.09	0.44**
Sample Size	**62**	**18**	**14**	**30**

Source: Q-CCIIT Fall 2012 Psychometric Field Test

[a] The item rating used in analysis is the mean across valid observation cycles (five minutes or longer) rounded to the nearest integer.

[b] Rated across the visit.

*p<.05; **p<.01; ***p<.001.

Table D.3. Test-Retest Correlations: Support for Cognitive Development Items

	Total	Infant	Toddler	FCC
Supporting object exploration[a]	0.61***	0.45*	0.64**	0.73***
Scaffolding problem solving[a]	0.70***	0.75***	0.65**	0.65***
Giving choices[b]	0.59***	0.41	0.57**	0.56***
Extending pretend play[b]	0.40***	0.54**	0.56**	0.11
Explicit teaching[b]	0.80***	0.89***	0.74***	0.75***
Supporting peer interaction/play[a]	0.70***	0.72***	0.74***	0.65***
Support for social problem solving[b]	0.48***	0.74***	0.77***	-0.10
Unique concepts	0.53***	0.47*	0.19	0.55***
Sample Size	**61-62**	**17-18**	**14**	**30**

Source: Q-CCIIT Fall 2012 Psychometric Field Test

[a] The item rating used in analysis is the mean across valid observation cycles (five minutes or longer) rounded to the nearest integer.

[b] Rated across the visit.

*p<.05; **p<.01; ***p<.001.

Table D.4. Test-Retest Mean Difference for Scale Scores Overall and by Subgroup

Scales	Total	Infant	Toddler	FCC
Support for Social-Emotional Development	0.51	0.37	0.68	0.51
Support for Language and Literacy Development	0.47	0.39	0.52	0.50
Support for Cognitive Development	0.48	0.34	0.39	0.61
Chaos	0.66	0.44	0.71	0.76
Extreme Areas of Concern	0.08	0.00	0.00	0.17
Sample Size	**61-62**	**18**	**14**	**29-30**

Source: Q-CCIIT Fall 2012 Psychometric Field Test

[a] The item rating used in analysis is the mean across valid observation cycles (five minutes or longer) rounded to the nearest integer.

[b] Rated across the visit.

*p<.05; **p<.01; ***p<.001.

Table D.5. Test-Retest Mean Difference: Support for Social-Emotional Development Items

	Total	Infant	Toddler	FCC
Responding contingently to distress[a]	0.96	1.38	1.05	0.66
Responding to social cues[a]	0.65	0.56	0.74	0.66
Responding to emotional cues[a]	0.55	0.45	0.69	0.55
Builds a positive relationship[a]	0.63	0.41	0.82	0.67
Supervises or joins in play and activities[b]	0.94	0.78	0.93	1.03
Responsive routines[b]	0.75	0.44	0.86	0.90
Classroom limits and management[b]	0.97	0.83	0.93	1.07
Sense of belonging[b]	0.66	0.61	0.86	0.60
Sample Size	**61-62**	**18**	**14**	**29-30**

Source: Q-CCIIT Fall 2012 Psychometric Field Test

[a] The item rating used in analysis is the mean across valid observation cycles (five minutes or longer) rounded to the nearest integer.

[b] Rated across the visit.

*p<.05; **p<.01; ***p<.001.

Table D.6. Test-Retest Mean Difference: Support for Language and Literacy Development Items

	Total	Infant	Toddler	FCC
Caregiver use of varied vocabulary[a]	0.72	0.67	0.78	0.72
Conversational turn-taking[a]	0.68	0.52	0.82	0.71
Use of questions[a]	0.49	0.31	0.54	0.57
Extending children's language use[a]	0.57	0.52	0.56	0.59
Engaging children in books[a]	0.38	0.35	0.39	0.39
Variety of words[a]	0.33	0.22	0.43	0.34
Variety of types of sentences[a]	0.33	0.25	0.44	0.32
Features of talk[b]	0.60	0.56	0.86	0.50
Talk about things not present[b]	1.02	0.78	1.14	1.10
Positive attitude toward books[b]	1.06	0.83	1.43	1.03
Sample Size	**62**	**18**	**14**	**30**

Source: Q-CCIIT Fall 2012 Psychometric Field Test

[a] The item rating used in analysis is the mean across valid observation cycles (five minutes or longer) rounded to the nearest integer.

[b] Rated across the visit.

*p<.05; **p<.01; ***p<.001.

Table D.7. Test-Retest Mean Difference: Support for Cognitive Development Items

	Total	Infant	Toddler	FCC
Supporting object exploration[a]	0.78	0.93	0.74	0.71
Scaffolding problem solving[a]	0.76	0.61	0.78	0.84
Giving choices[b]	1.00	1.06	1.00	0.97
Extending pretend play[b]	1.31	1.00	1.21	1.53
Explicit teaching[b]	0.44	0.29	0.57	0.47
Supporting peer interaction/play[a]	0.71	0.62	0.64	0.80
Support for social problem solving[b]	1.30	0.71	0.93	1.80
Unique concepts	0.81	0.83	0.93	0.73
Sample Size	**61-62**	**17-18**	**14**	**30**

Source: Q-CCIIT Fall 2012 Psychometric Field Test

[a] The item rating used in analysis is the mean across valid observation cycles (five minutes or longer) rounded to the nearest integer.

[b] Rated across the visit.

*p<.05; **p<.01; ***p<.001.

This page has been left blank for double-sided copying.

APPENDIX E

ITEM-LEVEL RELIABILITY

This page has been left blank for double-sided copying.

Table E.1. Support for Social-Emotional Development

Variable	Weighted Kappa
Responding to Distress	0.77
Responding to Social Cues	0.57
Responding to Emotional Cues	0.56
Building Positive Relationship	0.57
Supporting Peer Interaction	0.67
Support for Social Problem Solving	0.72
Responsive Routines	0.58
Classroom Limits/Management	0.66
Sense of Belonging	0.52
Sample Size	**52**

[a] Kappas are unweighted

[b] Items had only one rating selected

Table E.2. Support for Cognitive Development

Variable	Weighted Kappa
Support for Object Exploration	0.51
Scaffolding Problem Solving	0.56
Giving Choices	0.62
Explicit Teaching	0.58
Extending Pretend Play	0.65
Supervises/Joins in Play	0.60
Sum of all Concepts	0.74
Count of Unique Concepts	0.55
Sample Size	**52**

[a] Kappas are unweighted

[b] Items had only one rating selected

Table E.3. Support for Language Development

Variable	Weighted Kappa
Use of Varied Vocabulary	0.55
Conversational Turn Taking	0.41
Use of Questions	0.37
Extending Children's Language Use	0.48
Engaging Children in Books	0.60
Variety of Words	0.45
Variety of Sentences	0.45
Max Children in Bookshare	0.92
Sample Size	**52**

[a] Kappas are unweighted

[b] Items had only one rating selected

Table E.4. Environmental Items

Variable	Weighted Kappa
Schedule	0.34
Quiet Area	0.76
Indoor Space Organized	0.44
Support for Parents	0.78
Number of Parents	0.90
Gross Motor Activities	0.65[a]
Fine Motor Activities	0.92[a]
Arts and Crafts	0.92[a]
Music	0.81[a]
Sensory Play	0.83[a]
Dramatic Play	0.83[a]
Books/Story Telling	0.81[a]
Sample Size	**52**

[a] Kappas are unweighted

[b] Items had only one rating selected

Table E.5. Areas of Concern

Variable	Weighted Kappa
Physically Harsh	0.88[a]
Verbally Harsh	0.88
Restricts Children	0.54
Comm. Mismatch	0.65
Repeatedly Singles Out	0.59
Ignores Children	0.67
Children Unoccupied	0.60
Overwhelms Children	0.74
Children Stressed by Demands	0.74
Adult TV	1.00[b]
Children's Media	0.65
Poor Supervision	0.61
Unsafe Environment	0.84
General Health Provisions	0.90
Count of Extreme Concern	0.69
Level of Chaos	0.51
Sample Size	**52**

[a] Kappas are unweighted

[b] Items had only one rating selected

Table E.6. Areas of Extreme Concern

Variable	Weighted Kappa
Physically Harsh_Extreme	1.00[a]
Verbally Harsh_Exreme	1.00[b]
Restricts_Extreme	1.00[a]
Comm. Mismatch_Extreme	1.00[a]
Ignores Children_Extreme	-0.02[a]
Adult TV_Extreme	1.00[b]
Children's Media_Extreme	0.79[a]
Poor Supervision_Extreme	-.00[a]
Unsafe Environment_Extreme	1.00[b]
Gen. Health Provisions_Extreme	-0.00[a]
Sample Size	**52**

[a] Kappas are unweighted

[b] Items had only one rating selected

Table E.7. Collapsed Areas of Concern

Variable	Weighted Kappa
Physically Harsh_Collapsed	0.88[a]
Verbally Harsh_Collapsed	0.85[a]
Restricts_Collapsed	0.56[a]
Comm. Mismatch_Collapsed	0.79[a]
Repeatedly Singles Out_Collapsed	0.62[a]
Ignores Children_Collapsed	0.75[a]
Children Unoccupied_Collapsed	0.60[a]
Overwhelms Children_Collapsed	0.65[a]
Children Stressed_Collapsed	0.65[a]
Adult TV_Collapsed	1.00[b]
Children's Media_Collapsed	0.54[a]
Poor Supervision_Collapsed	0.60[a]
Unsafe Environment_Collapsed	0.81[a]
Gen. Health Provisions_Collapsed	0.94[a]
Sample Size	**52**

[a] Kappas are unweighted

[b] Items had only one rating selected

Table E.8. Types of Talk

Variable	Weighted Kappa
Parallel Language	0.29
Descriptive Talk	0.30
Songs and Poems	0.67
Explanations	0.65[a]
Reasoning	0.49[a]
Anticipatory	0.49
Decontextualized	0.61[a]
Sample Size	**52**

[a] Kappas are unweighted

[b] Items had only one rating selected

APPENDIX F

GENERALIZABILITY STUDY

This page has been left blank for double-sided copying.

MATHEMATICA
Policy Research

1100 1st Street, NE, 12th Floo1
Washington, DC 20002-4221
Telephone (202) 484-9220
Fax (202) 863-1763
www.mathematica-mpr.com

MEMORANDUM

TO: Amy Madigan, OPRE, ACF

FROM: Amang Sukasih, Xiaojing Lin, and Sameena Salvucci **DATE:** 9/10/2013

SUBJECT: Quality of Caregiver-Child Interactions for Infants and Toddlers (Q-CCIIT) Observation Tool: Reliability Estimates from the Perspective of Generalizability Theory

---

The Q-CCIIT observation tool is a new observational measure that was developed to assess the quality of child care settings, specifically the quality of caregiver-child interactions for infants and toddlers in non-parental care. The Q-CCIIT observational tool is appropriate for use across child care settings: center-based and family child care (FCC) settings, as well as single and mixed-age classrooms.

A psychometric field test was conducted in fall 2012 that involved observations using the Q-CCIIT observational tool with a sample of infant and toddler child care providers. We used data from the field test to conduct a study to assess the reliability of the Q-CCIIT measure using generalizability theory (Brennan 2000, 2001; Shavelson and Webb 1991). This memorandum presents the study methodology and results.

A. METHOD

1. Participants

The fall 2012 Q-CCIIT psychometric field test took place in 10 locations across the United States[1] and included a sample of 403 classrooms of which 73 percent were center-based (n=293) and 27 percent were in family child care (FCC) settings (n=110). Approximately 37 percent of the classrooms had both infants and toddlers (n=135), 34 percent had only toddlers (n=125), and 29 percent had only infants (n=108)[2].

Of the 2,158 toddlers and infants in the sampled classrooms, 87 percent of the children were in center-based settings (n=1,794) and 13 percent (n=364) were in FCC settings. Approximately 43 percent of these children were younger than 18 months, 47 percent were 18-36 months old,

[1] A separate memo summarizing the site selection approach was submitted to ACF on April 19, 2012.

[2] Information on the age of children in the classrooms was missing for 35 classrooms.

and approximately 10 percent had missing information for age. Also, 31 percent were dual language learners, and 15 percent had individualized family service plans.

2. Measures and Procedures

The Q-CCIIT instrument was designed to evaluate caregiver-child interactions that lead to more positive outcomes for infants and toddlers. Individual observations are organized around support for three key domains of children's development: social-emotional development, cognitive development, and language development with some practices that are rated across the full observation period.

The Q-CCIIT instrument was designed to be used by a trained observer, who rates the quality of caregiver interactions in each area such as support for social-emotional development (for example, responding contingently to distress and responding contingently to social cues) using seven-point rubrics.

Twenty one observers were trained to conduct classroom observations using the Q-CCIIT instrument. Each observer was assigned a subset of the 403 child care classrooms (observers were assigned anywhere from 1 to up to 34 classrooms each with 22 classrooms being the median number of classrooms assigned per observer). In other words, the classrooms were nested within observers. Each classroom observation lasted approximately two hours.

Within a classroom observation the observer conducted multiple timed *cycles* of observation for most items on the Q-CCIIT. The Q-CCIIT instrument includes two types of ratings: cycle ratings, and "across-the-visit" ratings. Cycle ratings were done each cycle (for example, there were six sets of ratings for items in sections A, B, and C of the Q-CCIIT) and "across-the-visit" ratings were done once for the entire 2-hour visit (for example, there was only one rating for items in section D).

Each cycle was 10 minutes long. In other words, the observer watched the caregiver in the classroom for 10 minutes and then rated the caregiver using the rubrics on the Q-CCIIT instrument, then watched the caregiver for a second cycle of 10 minutes and rated the caregiver again using the Q-CCIIT instrument, until the observer completed at least five cycles for each of the items that required a cycle rating. So within each classroom observation there were typically six scores for most items (for example, items in sections A, B, and C). However, not all items on the Q-CCIIT were rated in each cycle. For example, items in section D were rated only once across the 2-hour visit.

3. Dependent Variables

Dependent variables in our study were ratings of successful behavior in three specified areas of supporting children's development that were rated for each cycle: (1) support for social-emotional development, (2) support for language and literacy, and (3) support for cognitive development. These three areas included a set of items rated in each of at least five 10-minute cycles. Each of the items from Sections A, B, and C in the questionnaire (listed below) were scored on a rubric continuum ranging from 1 to 7 where the highest rating of 7 corresponded to a caregiver demonstrating the highest quality in supporting infants and toddlers. For example, for the first item under section A, "responding contingently to distress," a rating of 1 indicated that the caregiver ignored multiple distress signals, 3 indicated that the caregiver had a delayed response to distress signals, 5 indicated that the caregiver provided a soothing response to distress, and 7 indicated that the caregiver provided a calm, soothing response to distress while also balancing the needs of other children. Another set of items from section D of the instrument were rated once across the visit. These items cover all three of the areas that correspond to supporting children's development, as described above, and the three scales that are defined below show how each of the items in section D link to these areas.

4. Generalizability Study Design

Traditionally, classical test theory (CTT) is used to estimate the reliability of instruments (Allen, M.J. & Yen, Y.M. 2002) and continues to be used extensively. In CTT, variability in scores is partitioned into the variance due to the true scores and the variance due to error. A key assumption in CTT is that error is randomly distributed and comes from sources unrelated to true differences in the assessed trait (for example, the quality of child care). However, this approach has some limitations. A primary limitation is that there is only one undifferentiated error term in the traditional reliability coefficient formula, which does not allow us to address the multiple potential sources of error that exist in the Q-CCIIT instrument.

To address this limitation to the traditional CTT approach, we will use a different conceptual framework called Generalizability Theory (Brennan 2000; 2001; Shavelson and Webb 1991; Cronbach, Gleser, Nanda, and Rajaratnam 1972; Lord and Novick 1968) to assess the reliability of the Q-CCIIT measure. Generalizability theory (G theory) is composed of two phases: 1) a Generalizability study (G-study), which uses data to estimate the magnitude of the potential sources of error due to an observational measure like the Q-CCIIT and, 2) a Decision study (D-study) which applies the results of the G study to a range of possible future scenarios (for example, to determine the minimum number of Q-CCIIT observation cycles and observers needed to get adequate reliability) (Nunnally & Bernstein 1994). In Generalizability theory (G theory), a behavioral measurement score (for example, a child care quality observation score) is conceived of as a "sample from a *universe of admissible observations*" (Shavelson, R. J., & Webb, N. M., 1991, page 1). A measurement situation, like our Q-CCIIT field test, has characteristic features and "each characteristic feature is called a *facet* of a measurement. A universe of admissible observations, then is defined by all possible combinations of the levels of

the facets" (Shavelson, R. J., & Webb, N. M. (1991), page 1). G theory uses analysis of variance (ANOVA) procedures to obtain estimates of variance components for these different facets included in the analysis.

The universe of non-parental child care provider classrooms in our G-study is random since all such classrooms in the U.S. were not included in the study. The object of measurement in our G-study is the classroom, and the universe score is defined as the parameter mean score for infant and toddler classrooms. The facets in our G-study include observer, item, and cycle. For practical reasons, the Q-CCIIT field test data collection was designed with classrooms nested within observers (in other words, these variables were not completely crossed). Within each classroom an observer observed the classroom multiple times (cycles) for most items on the Q-CCIIT instrument. Therefore, our Q-CCIIT field test data collection design does not allow for using the fully crossed ANOVA models (cross-classified design with every observer observing the same classrooms, items, and occasions/cycles) used in many G-theory studies. The variance component associated with items represents the error associated with the sampling of the time of observations. Therefore, our G-study design allows for an estimation of variance components for item, observer by item interaction, and cycle by item interaction. The Q-CCIIT ANOVA model also assumes that any higher order interactions and the remaining unmeasurable variability are represented by a residual/error term. In addition, for the three specified areas for supporting children's development, different sets of items were used to measure each area (defined below).

Our G-study was conceptualized as a multi-faceted (as in the three facets introduced above), nested design. The dependent variable was the observational outcome, or score, on a 1 to 7 point Likert-type scale rubric for each of three support skill scales on each of up to seven cycles/occasions for each of the 403 classrooms. We will examine the variance component estimates for classrooms, observers, items, and cycles/occasions, and will decompose the estimate of variance components into the following terms: (1) classrooms nested within observers; (2) observers; (3) items; (4) occasions/cycles; (5) the interactions between observers and items, observers and cycles, items and cycles, items and classrooms nested within observers, and occasions and classrooms nested within observers; and (6) residual error.

Note that our Q-CCIIT G-study implemented a unique nested design where classrooms (the object of measurement) were nested within observers (facet). A more common G-study approach is to use a fully crossed design where the object of measurement (classroom) is observed by multiple observers, or if a fully crossed design cannot be conducted, then a nested design can be used where observers are nested within classrooms (not the other way around as in the Q-CCIIT design). The implication of our Q-CCIIT design where classrooms (the object of measurement) were nested within observers (facet) is that the interpretation of the variance due to classroom cannot be separated from the variance due to the interaction between classroom and observer. That is, the classroom nested within observer term in the analytic model actually consists of the classroom term plus the interaction between classroom and observer term.

Our study comprised two phases: a Generalizability Study (G-study) which used the Q-CCIIT field test data to estimate the magnitude of potential sources of error, and a Decision Study (D-study) which applied the results of the G-study to determine the minimum number of Q-CCIIT observation cycles and observers needed to get adequate reliability.

5. Comparison of Results Using Computational Approaches for Unbalanced and Balanced Designs

In our Q-CCIIT design/data, the number of classroom observations was not equal across observers (this is referred to as an unbalanced nested design). Also, the number of cycles was not always equal across classroom observations due to either missing values, or in some of our models for each of the three child development areas some items were not always observed in all seven cycles (some classrooms observations included items that were observed in only one cycle). As discussed in the G theory literature, when the sample sizes are unequal, several challenges may exist as discussed in Webb, Shavelson, and Haertel (2007, page 115):

> "Estimating variance components in unbalanced designs is not straightforward. Some or all methods have problems of computational complexity, distributional assumptions, biased estimation, require decisions that cannot be justified in the context of generalizability theory, or produce results that are inconclusive (Brennan, 2001)."

Our G-study (and D-study) results were developed using standard statistical software, SAS PROC VARCOMP (SAS 9.3), with REML method to estimate the variance components. SAS uses analysis of variance methods where the variance components are derived under the assumption of a balanced design (equal sample size for each facet). We used the original data with unequal number of cycles across classrooms observations. We did not drop, nor impute any missing data. We compared these variance component results for a few of the models to results from three different computational scenarios to determine if there were any significant differences in results. In other words, we compared results from our unbalanced design (scenario a below) with results from three alternative computational scenarios (scenarios b, c, and d below). Scenario results are available upon request.

a. Unbalanced Design using SAS PROC VARCOMP

In this scenario we used the original data with unequal number of cycles across classrooms observations. We did not drop, nor impute any missing data. Then, we used SAS PROC VARCOMP (SAS 9.3), with the REML method to estimate the variance components.

b. Unbalanced Design using urGENOVA

In this scenario we used the original data with unequal number of cycles across classroom observations. We did not drop, nor impute, any missing data. Then we used urGENOVA

software version 2.1 (Brennan 2001) to estimate variance components. This software was developed for estimating random effects variance components for both balanced and unbalanced designs for models that are complete (all interaction terms are included).

c. Balanced Design by Data Subsetting

In this scenario we dropped data for some cycles in order to make the number of cycles within each classroom observation included in the G study analysis the same. Then, we used SAS PROC VARCOMP (SAS 9.3), with the Restricted Maximum Likelihood Method (REML) to estimate the variance components.

d. Balanced Design by Imputation

In this scenario instead of dropping data for some cycles to make the number of cycles within each classroom observation the same, we imputed scores for missing cycles with the mean of scores within the classroom. Then, we used SAS PROC VARCOMP (SAS 9.3), with the REML method to estimate the variance components.

The G- and D-study results from the unbalanced designs using either the SAS PROC VARCOMP or the urGENOVA software yielded similar results confirming that the SAS procedure handled the calculation of variances appropriately when there is an unbalanced design. The variance results based on either subsetting or imputed data differed from those based on the unbalanced designs using SAS PROC VARCOMP (and urGENOVA), which indicates a potential bias in the estimates due to subsetting or imputed data. So we will be reporting results from the unbalanced design using the PROC VARCOMP scenario in the remainder of this memo because it keeps the original data intact avoiding potential biases in the results and uses standard commercial statistical software that is well tested.

B. RESULTS

Two sets of analyses were conducted: (1) G-study analyses, and (2) D-study analyses.

1. G-Study

In order to determine if the Q-CCITT was reliable, a series of multiple-facet, nested G-studies were conducted to compute the variance components using all the items in the measure (sections A, B, C, and D). The SAS PROC VARCOMP procedure estimates the contribution of each of the random effects to the variance of the dependent variable. Due to missing values of observations for some cycles within classrooms, the number of cycles is not equal across classrooms. The estimation method REML was used to allow unbalanced design for estimating effects from the general linear mixed model (Patterson & Thompson 1971). For the REML method we specified the maximum number of iterations to be 100.

Variance components and their proportions of contribution were calculated for seven different models. Models 1 through 4 did not include the ITEM term. Models 5 through 7 added the ITEM term in the models to see the variance proportion contribution due to ITEM. The model specifications for both the G and D studies are:

Model 1: Classrooms nested within observers; observers; and the residual term

Model 2: Classrooms nested within observers; cycles; observers; and the residual term

Model 3: Classrooms nested within observers; cycles; observers; cycle by observer interaction; and the residual term

Model 4: Classrooms nested within observers; cycles; observers; cycle by observer interaction; cycle by classroom nested within observer interaction; and the residual term

Model 5: Classrooms nested within observers; items; cycles; observers; and the residual term

Model 6: Classrooms nested within observers; items; cycles nested within observers; observers; and the residual term

Model 7: Classrooms nested within observers; observers; items; cycles; observer by item interaction; observer by cycle interaction; item by cycle interaction; item by classroom nested within observer interaction; cycle by classroom nested within observer interaction; and the residual term.

Results of the modeling are available upon request. Model 7 explained the greatest percentage of the variance and is described below. The model and statistics presented are given in the following mathematical formulas. Let indexes *o*, *c*, *i*, and *y*, respectively denote observer, class, item, and cycle. The variance associated with classroom scores can be decomposed into variance components as such:

$$\sigma^2_{X_{ocy}} = \sigma^2_{c(o)} + \sigma^2_o + \sigma^2_i + \sigma^2_y + \sigma^2_{oi} + \sigma^2_{oy} + \sigma^2_{iy} + \sigma^2_{ic(o)} + \sigma^2_{yc(o)} + \sigma^2_e.$$

In order to determine if the items from the three different Q-CCIIT scales were associated with different reliability estimates, separate G-studies were conducted according to the skill assessed. A series of three multi-facet, partially nested G-studies were conducted according to the three support skills assessed as defined below:

a. "Support for Social Emotional Development" which consisted of these items:

Item A1 (Responding Contingently to Distress)
Item A2 (Responding to Social Cues)
Item A3 (Responding to Emotional Cues)
Item A4 (Building Positive Relationship)
Item D8 (Supervises/Joins in Play)
Item D9 (Responsive Routines)
Item D10 (Classroom Management)
Item D11 (Sense of Belonging)

b. "**Support for Language and Literacy**" which consisted of these items:

Item C1 (Use of Varied Vocabulary)
Item C2 (Use of Questions)
Item C3 (Conversational Turn Taking)
Item C4 (Extending Language Use)
Item C5 (Engaging Children in Books)
Item C6 (Variety of Words)
Item C7 (Variety of Sentences)
Item D5 (Features of Talk)
Item D6 (Talk about Things not Present)
Item D7 (Positive Attitude toward Books)

c. "**Support for Cognitive Development**" which consisted of these items:

Constructed Variable CNTCONCPT (number of unique concepts)
Item A5 (Supporting Peer Interaction)
Item B1 (Object Exploration)
Item B2 (Scaffolding Problem Solving)
Item D1 (Giving Choices)
Item D2 (Extending Pretend Play)
Item D3 (Explicit Teaching)
Item D4 (Social Problem Solving)

Table 1 includes information for each of three support skills on the estimated variance components and the percentage of total variance explained by each facet and two-way interactions of Model 7.

Table 1: G-Study Model 7 Results: Variance Estimates and Percentage of Variance Attributable to Q-CCIITs Support Skills Explained by Each Facet and Key Interactions

	Support Skills					
	Social-Emotional Development		Language and Literacy		Cognitive Development	
Source of Variation	Estimated Variance	Percent of Variance	Estimated Variance	Percent of Variance	Estimated Variance	Percent of Variance
Classroom(Observer)	0.8755	38%	0.6274	25%	0.6141	25%
Observer	0.1407	6%	0.2722	11%	0.2118	9%
Item	0.0730	3%	0.3254	13%	0.0913	4%
Cycle	0.0008	<0.5%	0.0006	<0.5%	0.0000	0%
Observer x Item	0.0941	4%	0.1229	5%	0.1496	6%
Observer x Cycle	0.0105	<0.5%	0.0116	<0.5%	0.0197	1%
Item x Cycle	0.0001	<0.5%	0.0002	<0.5%	0.0020	<0.5%
Item x Class(Observer)	0.2701	12%	0.1799	7%	0.2556	10%
Cycle x Class(Observer)	0.4458	19%	0.4517	18%	0.3254	13%
Residual	0.3997	17%	0.5642	22%	0.8002	32%

Source: Q-CCIIT Fall 2012 Psychometric Field Test

Support for Social-Emotional Development. This was defined as the level of quality provided by caregivers in the setting to support the social-emotional development of children at the time of observation. The largest proportion of the variance for this scale was accounted for by the classroom nested within observer term, which is the object of measurement. Here, 38 percent of the variance in the provider's quality measure of support for children's social-emotional development was accounted for by individual classrooms, taking into account that observers were assigned to specific groups of classrooms for observations. The second largest contributor to the variance was the interaction term between the classrooms (nested within observers) and cycles (19 percent). The observer and item facets accounted for only about 6 percent and 3 percent, respectively, of the variance for this skill. These results indicate that classroom caregivers, nested within observer, varied in their ability to provide support for the social-emotional development of the children in their classrooms; it was the caregivers' quality level of the underlying skill, and not the rating style of the observer, that accounted for most of the variance.

Support for Language and Literacy Development. This was defined as the level of quality provided by caregivers in the setting to support the language and literacy development of children at the time of observation. The largest proportion of the variance for this skill was accounted for by the classroom nested within observer term (25 percent). The second largest contributor to the variance was the residual/error term. Here, 22 percent of the variance for the language and literacy development skill could not be accounted for by individual classrooms,

observers, items and cycles measurements, and their two-way interaction terms in Model 7. The observer facet accounted for only about 11 percent of the variance for this skill, while more variability in items (due to combining items from different sections of the instrument) contributed about 13 percent of the variance for this the language and literacy support skill. Nevertheless, these results indicate that classroom caregivers nested within observers varied in their ability to provide support for the language and literacy development of the children in their classrooms; it was the caregiver's level of this skill, and not the rating style of the observer, that accounted for most of the variance.

Support for Cognitive Development. This was as the level of quality provided by caregivers in the setting to support the cognitive development of children at the time of observation. The largest proportion of the variance for this skill was accounted for by the residual term (32 percent). The second largest contributor to the variance was the object of measurement, that is, classrooms nested within observers. Here, 25 percent of the variance for the cognitive development skill is accounted for by individual classrooms, taking into account that observers were assigned to specific groups of classrooms for observations. The observer facet accounts for only about 9 percent of the variance for this skill, and the item facet accounts for only about 4 percent. These results indicate that classroom caregivers nested within observer vary in their ability to provide support for the cognitive development of the children in their classrooms; it is the caregivers' level of this skill, and not the rating style of the observer, that accounts for most of the variance.

Despite that under our design the contribution of variance of classroom cannot be separated from that of observer, about one-third of the variability in measurement could be explained by differences among the child-care providers/classrooms (nested within observer) across the three skill scales in this study. The interaction terms between classroom (nested within observer) and cycle, as well as between classroom (nested within observer) and item, are not ignorable indicating that inconsistencies of the support may exist across cycles, or under different items. Finally, in the area of support for cognitive development, a significant amount of variability (about one-third of the total variance) was attributable to measurement error that cannot be explained by any particular identified facet of measurement, which in practical terms might mean that other potential sources of variance might include other facets not studied such as time of day or day of the week or the content of activities.

2. D-Study

A decision study (D study) is usually used to inform decisions about future approaches to measuring and collecting data from a given measure based on results from a G-study. In a D-study, decisions usually will be based on the mean of the variance over multiple observations rather than on a single observation. For example, a D study uses the coefficients generated in the G study (G-coefficient and the dependability index) to estimate what the reliability would be if different sizes/levels of facets were used. The reliability estimates in a D study for different levels of the facets are based on extrapolations using the coefficients generated in the G study.

In the Q-CCIIT design, however, it is difficult to conduct a D-study, not only due to the use of a nested design, but also a unique assignment of Q-CCIIT nested design where classrooms (the measurement object) were nested within observers (facet). Despite concerns about a D-study's limitations in this context, three different D-studies for each of the three skills were conducted in order to estimate how varying levels of cycles and observers might affect the reliability of the Q-CCIIT under Model 7. We calculated the statistics below for our D studies based on an assumption that the sample sizes are equal/balanced; that is, within each classroom n'_o number of observers observed n'_y cycles, and that the items used are those defined for each of the three scales (Section B1 includes the lists of items in each skill scale).

a. Relative error variance

The relative error is defined as the difference between an observed deviation score and its universe deviation score. For the Q-CCIIT under Model 7, it was defined and calculated as

$$\sigma_\delta^2 = \frac{\sigma_{ic(o)}^2}{n'_i n'_o} + \frac{\sigma_{yc(o)}^2}{n'_y n'_o} + \frac{\sigma_e^2}{n'_i n'_y n'_o},$$

and it can be interpreted as an estimate of the "relative" standard error of the measurement for a randomly selected classroom.

b. Absolute error variance

The absolute error is defined as the difference between an observed mean score and its universe score. For the Q-CCIIT it was defined and calculated as

$$\sigma_\Delta^2 = \frac{\sigma_o^2}{n'_o} + \frac{\sigma_i^2}{n'_i} + \frac{\sigma_y^2}{n'_y} + \frac{\sigma_{oi}^2}{n'_o n'_i} + \frac{\sigma_{oy}^2}{n'_o n'_y} + \frac{\sigma_{iy}^2}{n'_i n'_y} + \frac{\sigma_{ic(o)}^2}{n'_i n'_o} + \frac{\sigma_{yc(o)}^2}{n'_y n'_o} + \frac{\sigma_e^2}{n'_i n'_y n'_o},$$

and it can be interpreted as an estimate of the "absolute" standard error of the measurement for a randomly selected classroom.

c. The Generalizability Coefficient (G-coefficient)

The G-coefficient is defined as

$$E\rho^2 = \frac{\sigma^2_{c(o)}}{\sigma^2_{c(o)} + \sigma^2_{\delta}},$$

and it is the analogue of an internal consistency reliability coefficient in classical test theory.

d. The Dependability Index (Phi)

The dependability index is calculated as

$$\Phi = \frac{\boldsymbol{\sigma^2_{c(o)}}}{\boldsymbol{\sigma^2_{c(o)} + \sigma^2_{\Delta}}},$$

and has the same interpretation in terms of agreement as the G-coefficient.

The values for the estimates of the G-coefficient as well as the dependability index are given in Figures 1 through 6 for each of the three scales. We used results from our G-study, assuming that the number of items used is the same as in the current study, and that at least five cycles were observed by an observer within a classroom Under this design, the G coefficient and dependability index for each of the three support skill areas show an adequate level of reliability[3]; that is, they are mostly greater than 0.50. For example, the G-coefficient is 0.87 for support for social-emotional development, 0.84 for support for language and literacy, and 0.84 for support for cognitive development, when one observer was used to observe 5 cycles.

[3]Traditional cut-point interpretations for reliability estimates in G- and D-studies are that values around 0.2 are considered to correspond to low reliability, values around 0.5 can be considered to correspond to medium reliability, and values around 0.8 can be considered to correspond to high reliability (using Cohen's effect-size cut-offs).

Figure 1: D-Study Results under Model 7: G-Coefficient for Social-Emotional Development Skill, by Number of Observers and Number of Cycles

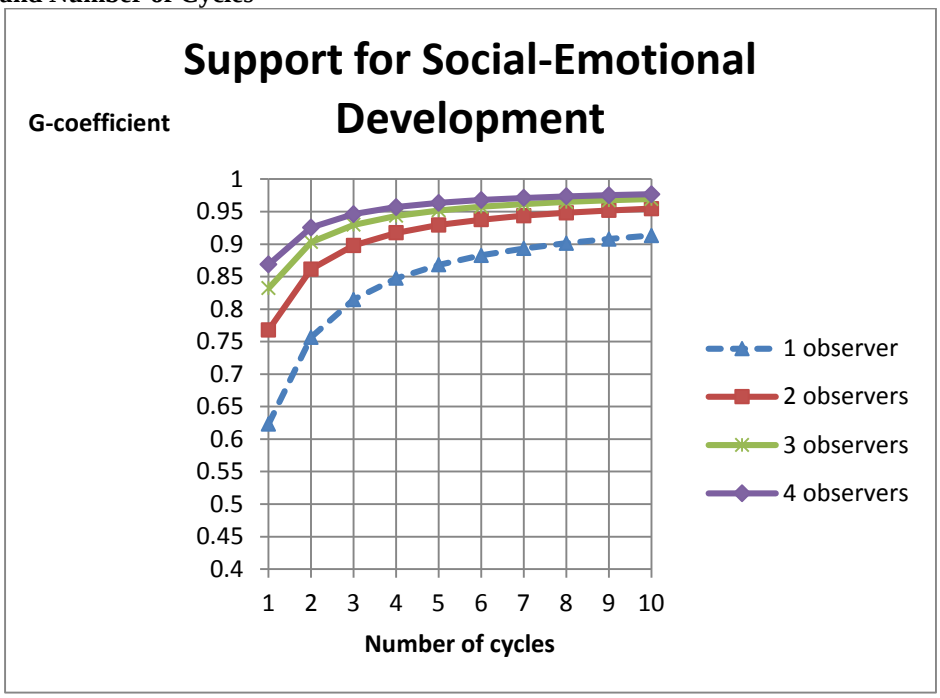


Figure 2: D-Study Results under Model 7: Dependability Index for Social-Emotional Development Skill, by Number of Observers and Number of Cycles

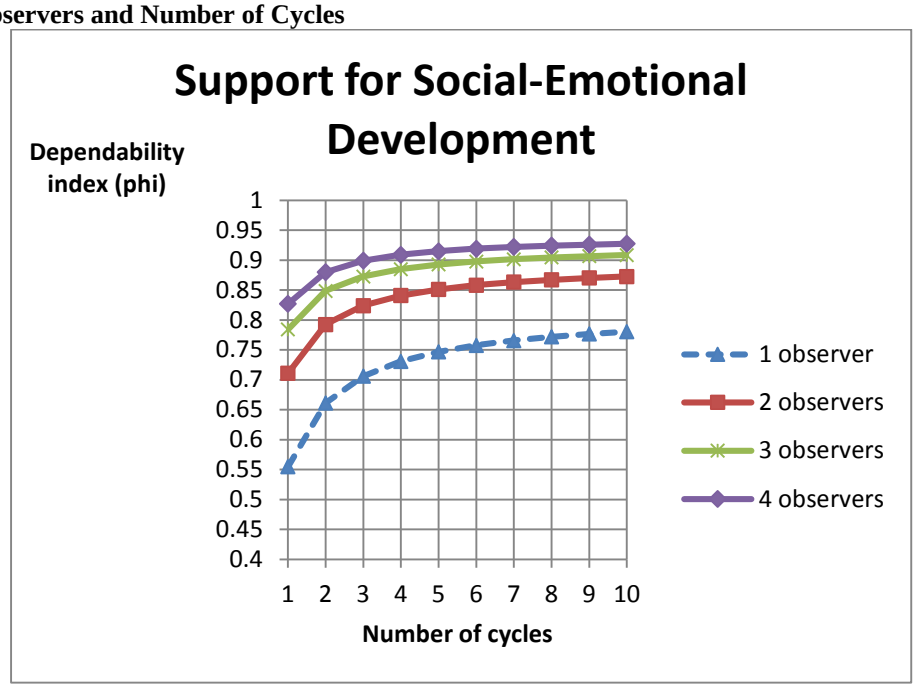


Figure 3: D-Study Results under Model 7: G-Coefficient for Language and Literacy Development Skill, by Number of Observers and Number of Cycles

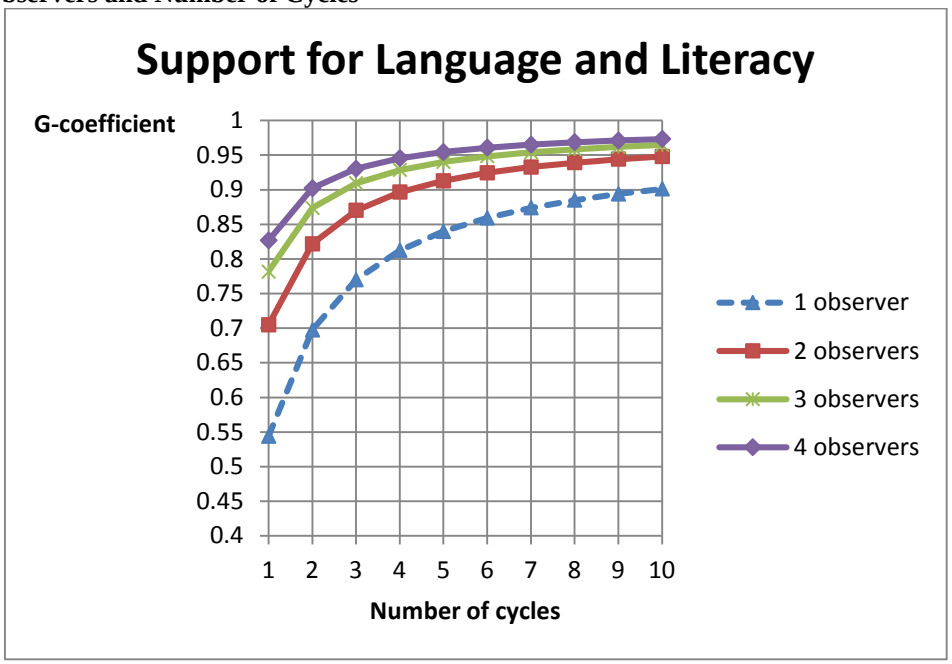


Figure 4: D-Study Results under Model 7: Dependability Index for Language and Literacy Development Skill, by Number of Observers and Number of Cycles

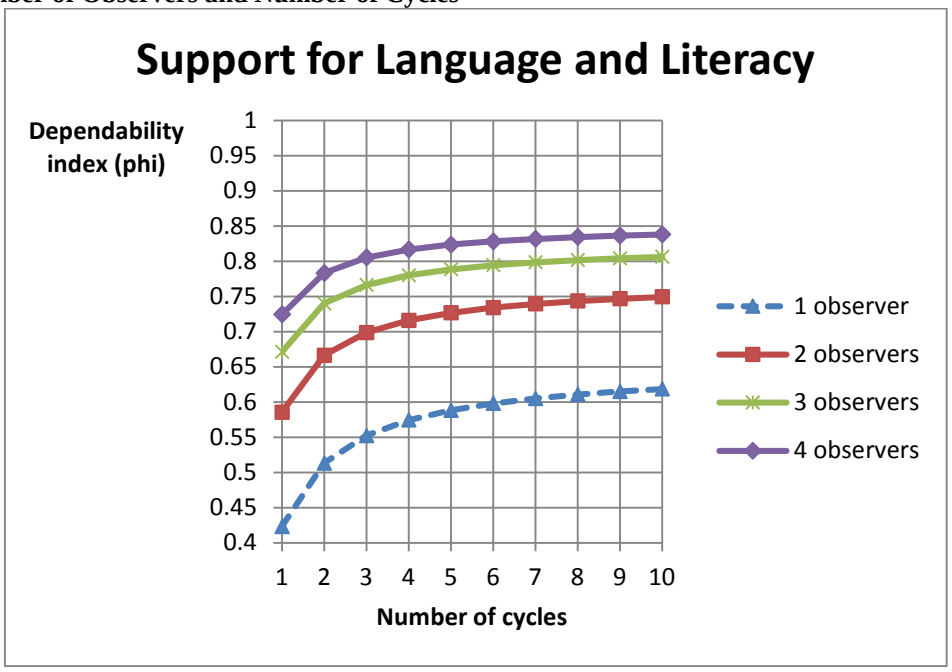


Figure 5: D-Study Results under Model 7: G-Coefficient for Cognitive Development Skill, by Number of Observers and Number of Cycles

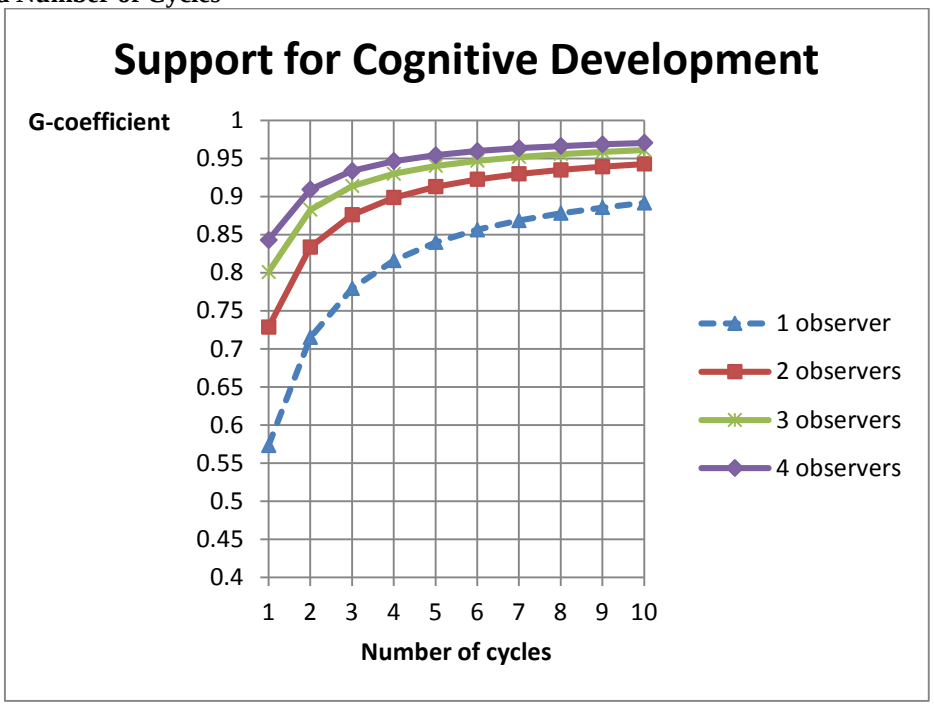
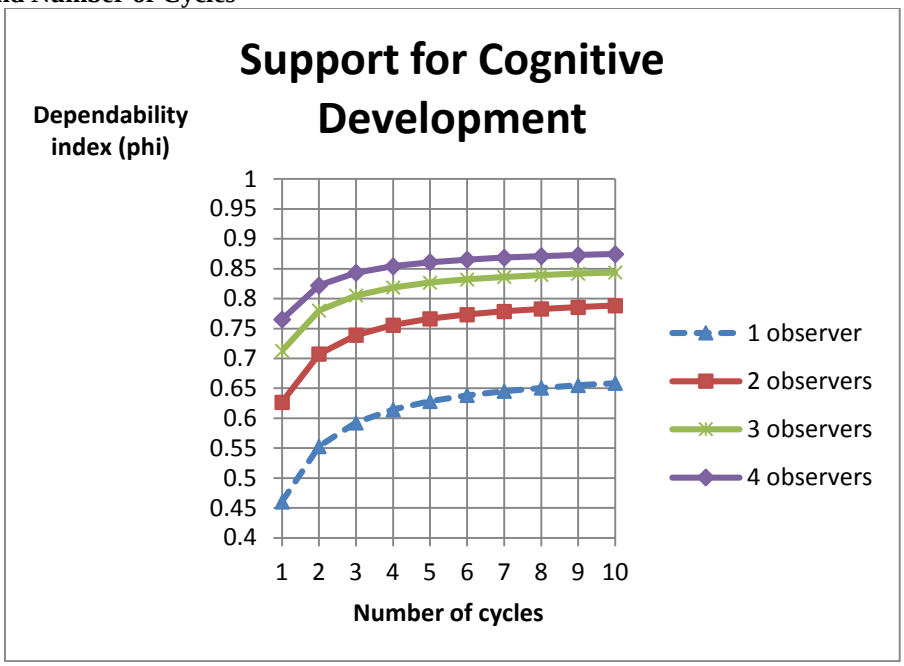


Figure 6: D-Study Results under Model 7: Dependability Index for Cognitive Development Skill, by Number of Observers and Number of Cycles



REFERENCES

Allen, M. J., & Yen, W. M. (2002*). Introduction to Measurement Theory. Long Grove, IL: Waveland Press.*

Brennan, R. L. (2000). "Performance assessments from the perspective of generalizability theory." *Applied Psychological Measurement*, 24, 339–353.

Brennan, R. L. (2001b). *Generalizability Theory*. Springer-Verlag.

Brennan, R. L. (2001b). Manual for urGENOVA. Iowa City, IA: Iowa Testing Programs, University of Iowa.Version 2.1

Lord, F. M., & Novick, M. R. (1968). Statistical theories of mental test scores. New Jersey: Addison-Wesley.

Patterson, H. D. and Thompson, R. (1971), "Recovery of Inter-Block Information When Block Sizes Are Unequal," *Biometrika*, 58, 545–554.

Nunnally, J., & Bernstein, I. (1994). *Psychometric Theory* (3rd ed., p. 736). McGraw-Hill Humanities/Social Sciences/Languages.

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability Theory: A Primer*. Newbury Park, CA: Sage.

Webb, N. M., Shavelson, R. J., & Haertel, E. H. (2007). Reliability Coefficients and Generalizability Theory. In C. R. Rao & S. Sinharay (Eds.), *Handbook of Statistics*, Vol. 26 (pp. 81-124).

APPENDIX G

SUBGROUP TABLES FOR CONVERGENT AND DISCRIMINANT VALIDITY

This page has been left blank for double-sided copying.

Table G.1. Correlations Between Q-CCIIT Scale Scores: Infants

	Support for Social-Emotional Development	Support for Language and Literacy Development	Support for Cognitive Development
Support for Social-Emotional Development	--		
Support for Language and Literacy Development	0.79***	--	
Support for Cognitive Development	0.76***	0.78***	--
Sample Size	**136**	**136**	**136**

Source: Q-CCIIT Fall 2012 Psychometric Field Test Data.

*p<.05; **p<.01; ***p<.001.

Table G.2. Correlations Between Q-CCIIT Scales and ITERS-R Subscales and Child-Adult Ratio: Infants

	Q-CCIIT Scales				
ITERS-R Subscales	Support for Social-Emotional Development	Support for Language and Literacy Development	Support for Cognitive Development	Concern extreme	Areas of Concern
Listening and Talking	0.48**	0.33	0.45**	0.14	-0.36*
Interaction Social	0.48**	0.35*	0.37*	-0.13	-0.41*
Activities	0.19	0.09	0.08	-0.12	0.02
Program Structure	0.51**	0.26	0.28	-0.30	-0.31
Space and Furnishings	0.23	0.26	0.23	-0.09	0.04
Personal Care	0.19	0.12	-0.04	-0.05	0.06
ITERS-R Total	0.42*	0.29	0.28	-0.12	-0.20
Child/Adult Ratio	-0.33	-0.13	-0.33	-0.18	0.34
Sample Size	**32**	**32**	**32**	**31**	**31**

Source: Q-CCIIT Fall 2012 Field Test Data.

*p<.05; **p<.01; ***p<.001.

ITERS-R = Infant-Toddler Environment Rating Scale-Revised.

Table G.3. Correlations Between Q-CCIIT Scales and the ORCE Scales: Infants

	Q-CCIIT Scales				
ORCE Scales	Support for Social-Emotional Development	Support for Language and Literacy Development	Support for Cognitive Development	Concern Extreme	Areas of Concern
Overall Qualitative Rating	0.57***	0.38*	0.48**	-0.17	-0.51**
Sensitivity/Responsiveness to Distress [a]	0.15	0.05	0.19	-0.10	-0.09
Sensitivity/Responsiveness to Non-Distress	0.63***	0.50**	0.58***	-0.43**	-0.58***
Lack of Intrusiveness	0.49**	0.38*	0.38*	-0.61***	-0.42**
Lack of Detachment/Disengagement	0.32	0.22	0.20	-0.03	-0.29
Stimulation of Cognitive Development	0.61***	0.42**	0.56***	-0.23	-0.50**
Positive Regard for the Child	0.34*	0.23	0.37*	0.12	-0.34*
Lack of Negative Regard for the Child	-0.01	-0.08	0.06	0.04	-0.00
Flatness of Affect	0.25	0.08	0.07	0.01	-0.22
Fostering Exploration[b]	.	.	.	.	.
Positive Rating	0.62***	0.45**	0.60***	-0.24	-0.56***
Lack of Negative Rating	0.36*	0.21	0.20	-0.13	-0.32
Language Stimulation	0.24	0.05	0.28	-0.07	-0.24
Positive Behavior Toward Child	-0.06	-0.10	-0.00	0.30	0.00
Negative Behavior Toward Child	-0.49**	-0.46**	-0.56***	0.52***	0.37*
Sample Size	**32-37**	**32-37**	**32-37**	**32-37**	**32-37**

Source: Q-CCIIT Fall 2012 Field Test Data.

*p<.05; **p<.01; ***p<.001.

ORCE = Observational Ratings of the Caregiving Environment.

Table G.4. Correlations Between Q-CCIIT Scale Scores, Toddlers

	Support for Social-Emotional Development	Support for Language and Literacy Development	Support for Cognitive Development
Support for Social-Emotional Development	--		
Support for Language and Literacy Development	0.80***	--	
Support for Cognitive Development	0.72***	0.72***	--
Sample Size	**154**	**154**	**154**

Source: Q-CCIIT Fall 2012 Psychometric Field Test Data.

*p<.05; **p<.01; ***p<.001.

Table G.5. Correlations Between Q-CCIIT Scales and ITERS-R Subscales and Child-Adult Ratio, Toddlers

	Q-CCIIT Scales				
ITERS-R Subscales	Support for Social-Emotional Development	Support for Language and Literacy Development	Support for Cognitive Development	Concern Extreme	Areas of Concern
Listening and Talking	0.66***	0.45**	0.34	.	-0.48**
Interaction Social	0.51**	0.14	0.19	.	-0.51**
Activities	0.57***	0.30	0.35*	.	-0.39*
Program Structure	0.56***	0.45*	0.11	.	-0.38*
Space and Furnishings	0.33	0.43*	0.17	.	-0.00
Personal Care	0.42*	0.29	0.17	.	-0.17
ITERS-R Total	0.66***	0.40*	0.30	.	-0.45*
Child/Adult Ratio	-0.29	-0.20	-0.22	.	0.33
Sample Size	**31-32**	**31-32**	**31-32**		**31-32**

Source: Q-CCIIT Fall 2012 Psychometric Field Test Data.

*p<.05; **p<.01; ***p<.001.

ITERS-R = Infant-Toddler Environment Rating Scale-Revised.

Table G.6. Correlations Between Q-CCIIT Scales and the ORCE Scales, Toddlers

	Q-CCIIT Scales				
ORCE Scales	Support for Social-Emotional Development	Support for Language and Literacy Development	Support for Cognitive Development	Concern Extreme	Areas of Concern
Overall Qualitative Rating	0.57***	0.57***	0.41**	-0.27	-0.62***
Sensitivity/Responsiveness to Distress[a]	0.05	0.10	0.21	0.00	-0.25
Sensitivity/Responsiveness to Non-Distress	0.58***	0.57***	0.45**	-0.46**	-0.64***
Lack of Intrusiveness	0.41**	0.44**	0.29	-0.41**	-0.62***
Lack of Detachment/Disengagement	0.38*	0.34*	0.21	-0.17	-0.53***
Stimulation of Cognitive Development	0.47**	0.47**	0.39*	-0.31*	-0.38*
Positive Regard for the Child	0.48**	0.56***	0.43**	-0.31*	-0.51***
Lack of Negative Regard for the Child	0.03	0.12	0.18	0.07	-0.06
Flatness of Affect	0.33*	0.35*	0.21	0.01	-0.38*
Fostering Exploration[b]	0.43	0.50*	0.26	.	-0.44
Positive Rating	0.58***	0.59***	0.45**	-0.36*	-0.58***
Lack of Negative Rating	0.44**	0.44**	0.28	-0.17	-0.58***
Language Stimulation	0.53***	0.50***	0.38*	-0.19	-0.41**
Positive Behavior Toward Child	0.38*	0.37*	0.35*	-0.07	-0.23
Negative Behavior Toward Child	0.04	-0.02	-0.07	0.02	-0.00
Sample Size	**19-41**	**19-41**	**19-41**	**19-41**	**19-41**

Source: Q-CCIIT Fall 2012 Psychometric Field Test Data.

*p<.05; **p<.01; ***p<.001.

ORCE = Observational Ratings of the Caregiving Environment.

Table G.7. Correlations Between Q-CCIIT Scale Scores, FCCs

	Support for Social-Emotional Development	Support for Language and Literacy Development	Support for Cognitive Development
Support for Social-Emotional Development	--		
Support for Language and Literacy Development	0.81***	--	
Support for Cognitive Development	0.80***	0.79***	--
Sample Size	**110**		

Source: Q-CCIIT Fall 2012 Psychometric Field Test Data.

*p<.05; **p<.01; ***p<.001.

Table G.8. Correlations Between Q-CCIIT Scales and the ORCE Scales, FCCs

	Q-CCIIT Scales				
ORCE Scales	Support for Social-Emotional Development	Support for Language and Literacy Development	Support for Cognitive Development	Concern Extreme	Areas of Concern
Overall Qualitative Rating	0.45**	0.26	0.39*	-0.27	-0.38*
Sensitivity/Responsiveness to Distress [a]	0.09	-0.34	-0.05	-0.07	-0.34
Sensitivity/Responsiveness to Non-Distress	0.46**	0.27	0.35*	-0.37*	-0.33*
Lack of Intrusiveness	0.36*	0.18	0.30	-0.37*	-0.47**
Lack of Detachment/Disengagement	0.42**	0.29	0.41**	-0.09	-0.32*
Stimulation of Cognitive Development	0.32*	0.18	0.16	-0.22	-0.21
Positive Regard for the Child	0.39*	0.21	0.31*	-0.37*	-0.51***
Lack of Negative Regard for the Child	0.34*	0.24	0.23	-0.38*	-0.45**
Flatness of Affect	0.25	0.22	0.29	-0.02	-0.21
Fostering Exploration[b]	0.51*	0.23	0.43	-0.33	-0.41
Positive Rating	0.42**	0.23	0.33*	-0.33*	-0.35*
Lack of Negative Rating	0.42**	0.28	0.40**	-0.21	-0.41**
Language Stimulation	0.24	0.21	0.11	-0.06	-0.20
Positive Behavior Toward Child	0.24	0.11	-0.01	-0.29	-0.30
Negative Behavior Toward Child	-0.52***	-0.43**	-0.41**	0.33*	0.28
Sample Size	**20-41**	**20-41**	**20-41**	**20-41**	**20-41**

Source: Q-CCIIT Fall 2012 Field Test Data.-

*p<.05; **p<.01; ***p<.001.

ORCE = Observational Ratings of the Caregiving Environment.

APPENDIX H

THE Q-CCIIT ENVIRONMENTAL ITEMS

This page has been left blank for double-sided copying.

The Q-CCIIT included four items about positive aspects of the environment beyond what is captured in the interaction scales and areas of concern: (1) schedule balances types of activities, (2) quiet area is available, (3) caregiving space is organized, and (4) caregivers are supportive of parents (rated based on actual observations with parents; if no parents observed, rated NA). Our experts recommended that we add items addressing these areas. The Q-CCIIT also included an additional categorical item—Count of Activities—created from a number of different types of activities: gross motor, fine motor, arts and crafts, music, sensory play, dramatic play, and books/storytelling (other than Goodnight Gorilla). Observers checked all types of activities available to children, even if no children engaged in those activities; however, each activity could be assigned to only one category. Based on the distribution, we created a 1 to 4 score: 1—if one to three different activities were available; 2—if four or five activities were available; 3—if six activities were available; and 4—if more than six activities were available. In this section, we present analyses of the four environmental items, with and without the additional Count of Activities item.

A. Item Descriptives

Of the four environment items included in the Q-CCIIT, "caregivers are supportive of parents" had the highest mean rating (3.67 out of 4.00) across the entire sample, followed by "caregiving space is organized" (3.61), "schedule balances types of activities" (3.37), and "quiet area is available" (2.70) (Table H.1). "Caregivers are supportive of parents" also had the highest percentage of missing data, since observers were able to rate that item only if they observed the caregiver interacting with a parent. The categorical Count of Activities item had a mean of 2.30 with no missing data, indicating that the average classroom had more than three activities available.

Table H.1. Mean of Q-CCIIT Environmental Scale Items for the Overall Sample

			Center-Based							
	Overall Total		Total		Infants		Toddlers		FCC Total	
	Mean (SD)	% Missing	Mean (SD)	% Missing	Mean (SD)	% Missing	Mean (SD)	% Missing	Mean (SD)	% Missing
Caregivers Are Supportive of Parents	3.67 (0.72)	34.50	3.75 (0.62)	29.66	3.83 (0.52)	30.15	3.67 (0.69)	29.22	3.40 (0.95)	47.27
Caregiving Space Is Organized	3.61 (0.71)	2.50	3.76 (0.55)	2.07	3.70 (0.62)	2.21	3.81 (0.47)	1.95	3.23 (0.93)	3.64
Schedule Balances Types of Activities	3.37 (0.85)	0.75	3.44 (0.80)	1.03	3.34 (0.88)	1.47	3.54 (0.73)	0.65	3.19 (0.92)	0.00
Quiet Area Is Available	2.70 (1.38)	0.50	2.94 (1.32)	0.34	2.87 (1.33)	0.00	3.00 (1.30)	0.65	2.06 (1.34)	0.91
Count of Activities	2.30 (0.97)	0.00	2.31 (0.89)	0.00	1.87 (0.69)	0.00	2.69 (0.87)	0.00	2.29 (0.81)	0.00
Sample Size	**400**		**290**		**136**		**154**		**110**	

Source: Q-CCIIT Fall 2012 Psychometric Field Test.

Range of all items was 1.00 – 4.00.

The Overall Environment Score (averaged across the four base environment items) had a mean of 3.28 across the entire sample, with a minimum of 1.00 and a maximum of 4.00 (Table H.2). Centers averaged a higher Overall Environment Score (3.43) than FCCs (2.89), and toddlers had a higher mean (3.48) than infants (3.37). When the Count of Activities item was included in the Overall Environment Score, mean scores decreased, both overall and for centers.

Missing data were not randomly distributed. Almost half of the FCCs (47.2 percent) were missing a rating for the item "caregivers are supportive of parents," compared with less than one-third of center-based classrooms that were missing this rating. It may be that children in FCCs arrive much earlier in the morning or, alternatively, that caregivers who are more supportive of parents asked that observers arrive later so they had time to support parents as they transitioned children into care. Although FCCs had a lower mean score on this item than center-based classrooms, the FCC mean was lower for all of the environment items except count of activities.

Table H.2. Mean and Range of Q-CCIIT Overall Environment Score for the Overall Sample, and by Child Age and Program Type

			Center-Based			
		Overall Total	Total	Infants	Toddlers	FCC Total
4 base items only	Mean	3.28 (0.70)	3.43 (0.60)	3.37 (0.65)	3.48 (0.56)	2.89 (0.78)
	Minimum	1.00	1.67	1.67	1.67	1.00
	Maximum	4.00	4.00	4.00	4.00	4.00
4 base items plus Count of Activities	Mean	3.07 (0.61)	3.19 (0.53)	3.05 (0.55)	3.31 (0.49)	2.76 (0.68)
	Minimum	1.00	1.60	1.60	1.75	1.00
	Maximum	4.00	4.00	4.00	4.00	4.00
Sample Size		**400**	**290**	**136**	**154**	**110**

Source: Q-CCIIT Fall 2012 Psychometric Field Test.

B. Assessing Reliability

We computed the coefficient alpha for the Q-CCIIT Overall Environment Score to assess internal consistency reliability (Table H.3). For the full sample for the Overall Environment Score, the internal consistency reliability estimates were weak but in an acceptable range (0.63). When broken out by child age and program type, however, the reliability estimates no longer fall within an acceptable range. Analysis of the item-to-total correlations suggested that the quiet area item was not contributing much to the scale. We examined the reliability with and without the quiet area item and with and without diverse activities. Reliability was similar with and without the quiet area item. The reliability estimate for the full sample was slightly weaker when the Count of Activities item was included (0.59).

Table H.3. Internal Reliability Estimates (Cronbach Alpha) of the Q-CCIIT Overall Environment Score for the Overall Sample, and by Child Age and Program Type

		Center-Based			
	Overall Total	Total	Infants	Toddlers	FCC Total
4 base items only	0.63	0.55	0.49	0.62	0.67
4 base items plus Count of Activities	0.59	0.49	0.50	0.52	0.73

Source: Q-CCIIT Fall 2012 Psychometric Field Test.

C. Assessing Convergent Validity

We examined associations with the concurrent observations on the Infant/Toddler Environment Rating Scale-Revised (ITERS-R) or Family Child Care Environment Rating Scale-Revised (FCCERS-R), depending on setting type. We also examined associations with the concurrent observations using the Observational Record of the Caregiving Environment (ORCE) in all settings. We examined concurrent validity by program type (centers and FCCs) and child age (infant and toddler classrooms). Since most of the FCCs include mixed-age groups, the analyses by child age are limited to infant and toddler classrooms in centers.

1. Convergent Validity with the ITERS-R and FCCERS-R Space and Furnishings Items

Across both center-based classrooms (ITERS-R) and FCCs (FCCERS-R), we found weak evidence of convergent validity for the Q-CCIIT Overall Environment Score (Table H.4). The ITERS-R space and furnishings score was correlated with the Q-CCIIT Overall Environment Score in center-based classrooms ($r = 0.30$), driven by a stronger correlation with toddler classrooms ($r = 0.54$). For infant classrooms, the correlation was weaker and non-significant ($r = 0.15$). When the Count of Activities item was included in the Overall Environment Score, the correlation for center-based classes improved slightly ($r = 0.31$); subgroup analyses indicated a slightly weaker relationship for toddlers ($r = 0.51$) and a slightly stronger—though still non-significant—correlation for infants ($r = 0.18$). The FCCERS-R space and furnishings score also was correlated with the Q-CCIIT Overall Environment Score in FCCs ($r = 0.29$); when the Count of Activities item was included in the Overall Environment Score, the correlation improved slightly ($r = 0.35$). The Q-CCIIT Overall Environment Score also was correlated with other FCCERS-R scales, including listening and talking ($r = 0.39$), interaction/social mean ($r = 0.33$), program structure ($r =0.30$), and activities ($r = 0.28$).

Table H.4. Correlations with Q-CCIIT Overall Environment Score for the Overall Sample, and by Child Age and Program Type

		Overall Total	Center-Based			FCC
			Total	Infants	Toddlers	
4 base items only	ITERS-R/FCCERS-R Space and Furnishings Mean	—	0.30*	0.15	0.54**	0.29*
	ORCE Overall Qualitative Rating	0.34***	0.43***	0.35*	0.54***	0.27
4 base items plus Count of Activities	ITERS-R/FCCERS-R Space and Furnishings Mean	—	0.31*	0.18	0.51**	0.35*
	ORCE Overall Qualitative Rating	0.32***	0.37***	0.34*	0.49**	0.30
Sample Size		**119**	**64−78**	**32−37**	**32−41**	**41−49**

Source: Q-CCIIT Fall 2012 Psychometric Field Test.

*p<.05; **p<.01; ***p<.001.

ITERS-R = Infant/Toddler Environment Rating Scale-Revised.
FCCERS-R = Family-Child Care Environment Rating Scale-Revised.
ORCE = Observational Ratings of the Caregiving Environment.

2. Relationship with the ORCE Overall Qualitative Rating, the ITERS-R Total Score, and FCCERS-R Total Score

The bivariate correlations between the ORCE Overall Qualitative Rating and the Q-CCIIT Overall Environment Score (Table 4) fell in the moderate range for the overall sample (r =0.34) as well as the center-based sample (r =0.43). No relationship was detected for FCCs. When the Count of Activities item was included in the Overall Environment Score, the correlation decreased for the overall sample (r =0.32) as well as for center-based classrooms, infants, and toddlers. The correlation for FCCs remained non-significant.

The correlations with the ORCE were greater than expected, given that the ORCE Overall Qualitative Rating focuses on relationships. After controlling for the Q-CCIIT's positive support scales (particularly Support for Social-Emotional Development) and Areas of Concern, the ORCE no longer showed a relationship with the Q-CCIIT Overall Environment Score (Table H.5). We had similar findings for the ITERS-R and FCCERS-R total scores. The environment items did not explain additional variance in the ORCE, ITERS-R, or FCCERS-R beyond what is explained by the other Q-CCIIT scales.

Table H.5. Relationship Between ORCE Qualitative Items and Q-CCIIT Environmental Scale, Controlling for Other Q-CCIIT Scales: Weighted OLS Results

	ORCE Overall Qualitative Rating	ITERS-R Total	FCCERS-R Total
Support for Social-Emotional Development	.346*	.539**	-.148
Support for Language and Literacy Development	-.121	. .048	-.103
Support for Cognitive Development	.109	-.110	.507*
Areas of Concern	-.220*	-.094	-.288+
Environment Items (without Count of Activities)	.059	.048	.083
R^2	**.294**	**.330**	**.326**

Source: Q-CCIIT Fall 2012 Psychometric Field Test.

Note: Model results reported as Beta coefficients.

+p<.10; *p<.05. **p<.01; ***p≤.001.

D. Summary

The Q-CCIIT Overall Environment Score has low reliability, a high percentage of missing data, and very limited evidence of convergent validity with other observational measures of the environment. No relation was found with the environmental scale after controlling for the other Q-CCIIT scales. Ultimately, our analyses do not provide support for the use of this set of items.

This page has been left blank for double-sided copying.


Q-CCIIT