PennState

CLEARINGHOUSE
FOR MILITARY FAMILY READINESS

Early Childhood Education (ECE) Project

Final Report on the Role that Military-Supported Early Childhood Education Plays in Promoting Child Development and School Readiness

Jennifer K. Karre, Ph.D.

Ryan P. Chesnut, Ph.D.

Molly Burns, M.A.

Katie E. Davenport, Ph.D.

Daniel F. Perkins, Ph.D.

27 January 2022

This material is the result of partnership funded by the Department of Defense between the Office of the Deputy Assistant Secretary of Defense for Military Community and Family Policy and the USDA's National Institute of Food and Agriculture through a grant/cooperative agreement with Penn State University.

Table of Contents

Executive Summary

In collaboration with the Office of the Secretary of Defense – Office of Military Community and Family Policy, the Clearinghouse for Military Family Readiness at Penn State (Clearinghouse) implemented the Early Childhood Education (ECE) Project, which was an evaluation of the Child Development Center (CDC) program. This effort was completed to better understand the role military-supported early childhood education plays in the promotion of children's development and school readiness and parental absenteeism from work. Data were collected from 20 military Child Development Centers (CDCs) and civilian child care centers. Data were collected from parents, direct-care staff, and directors, and by independent observers.

The evaluation was outcome-focused, which means the evaluation was intended to assess the degree to which participants achieved the intended outcomes of the program. The following four questions guided the evaluation:

- Does the CDC program improve child development outcomes among program participants?
- Is the CDC program more or less effective across different groups of participants (e.g., children who have direct-care staff with more or less training)?
- Do children attending military CDC programming differ from military children in comparable civilian programs? If so, how do they differ?
- How do children in CDC programming differ from normative comparison groups?

The evaluation design was quasi-experimental; it compared military children in CDCs and military children in civilian child care centers. Child development was intended to be assessed over the course of a year at 3-month intervals for a total of five timepoints (i.e., five waves). Parent outcome, child psychological well-being, process, demographic, and moderator variable measures were intended to be collected at the baseline, 6 months, and 1-year data-collection timepoints. To minimize burden, only the child development measures were originally scheduled to occur at all five waves.

Findings from analyses of the data indicated the following:

- The CDCs had higher classroom quality scores than the civilian centers.
- Direct-care staff education and perceived leadership support from directors and command were associated with classroom quality.
- Children in CDCs had better scores than the normed sample on several well-being indicators, whereas children in the civilian centers were not different from the normed sample on the well-being indicators.
- The results for the comparisons to the normed sample for the direct-care staff report of child development were mixed. An example follows:
    - At the Fall assessment, for five of the six development domains, fewer (i.e., a lower percent) 2-year-olds were not meeting expectations than in the normed sample; this is an indicator of greater development for the children in the CDCs.
    - However, also at the Fall assessment, for four of the six development domains, more (i.e., a higher percent) 1-year-olds were not meeting expectations than in

the normed sample; this is an indicator of less development for the children in the CDCs.

- Child well-being was related to attending a CDC, as opposed to a civilian center, and attending a CDC with more staff-perceived director support.
- More advanced child development was related to attending a CDC, as opposed to a civilian center; attending a CDC with more staff-perceived command support; and the child being in a higher-quality classroom.
- Parents in the CDCs reported missing fewer days of work due to child care arrangements compared to parents in the civilian centers.
- Parents reported missing fewer days of work when child care was available when the parents had to work late.

The evaluation, however, included several limitations that must be considered when interpreting the results. The limitations include the following:

- The COVID-19 pandemic impacted this evaluation substantially.
- Attrition and missing data from the parent reports and missing data from the direct-care staff reports of child development and classroom environment were present.
- Evidence of differential attrition was found; therefore, differences that appear in later waves must be understood in that context.
- The sample size for children in the civilian centers was small; therefore, additional differences may have been present that could not be detected due to the sample size.

Based on what was learned over the course of this evaluation, the evaluation team puts forth the following recommendations, which are consistent with the conclusions and recommendations in the National Academies of Sciences, Engineering, and Medicine's (2019) report on the military family readiness system. The recommendations fall into three categories: (1) continuous quality improvement, (2) leveraging data and existing expertise, and (3) organizational climate.

Recommendation 1: In addition to the trainings currently offered to direct-care staff, implement professional development activities informed by this and future data collection (e.g., the specific classroom-quality indicators that were identified, in this evaluation, as areas for improvement).

Recommendation 2: Leverage the standards of state and national early childhood education organizations (e.g., National Association for the Education of Young Children education standards) to inform decision-making.

Recommendation 3: Monitor organizational climate and use evidence-informed practices and information from this evaluation and future data-collection efforts to influence climate.

Additional Consideration: Given the unanticipated challenges encountered during this evaluation (i.e., the COVID-19 pandemic) and the implementation of a new curriculum, additional evaluations would provide further, and more complete, data regarding military supported early childhood education. An implementation evaluation would provide important information about how the new curriculum is operating on the ground, and an outcome evaluation would provide data that was disrupted by the COVID-19 pandemic.

Introduction

In collaboration with the Office of the Secretary of Defense (OSD) – Office of Military Community and Family Policy, the Clearinghouse for Military Family Readiness at Penn State (Clearinghouse) implemented the Early Childhood Education (ECE) Project, which was an evaluation of the Child Development Center (CDC) program. The CDC program consists of center-based child care for children 6 weeks to 5 years old and is located on military installations. CDC programs are inspected and certified by the Department of Defense (DoD) and each of the Services. In addition, CDCs are required to be accredited by a national accrediting body (e.g., National Association for the Education of Young Children [NAEYC]). CDCs offer full- and part-time child care that is consistent with typical working hours or mission requirements. Some CDCs offer extended care (e.g., before or after typical hours).

The goal of the ECE Project was to understand the role military-supported early childhood education plays in the promotion of children's development and school readiness and parental absenteeism from work. In order to do this, data were collected from parents, direct-care staff, CDC directors, and independent observers. Data were scheduled to be collected at five timepoints over 1 year. Three principles were followed in the development and execution of this evaluation: (1) use scientifically-sound evaluation methods, (2) minimize burden on CDC directors and staff; and (3) ensure confidentiality.

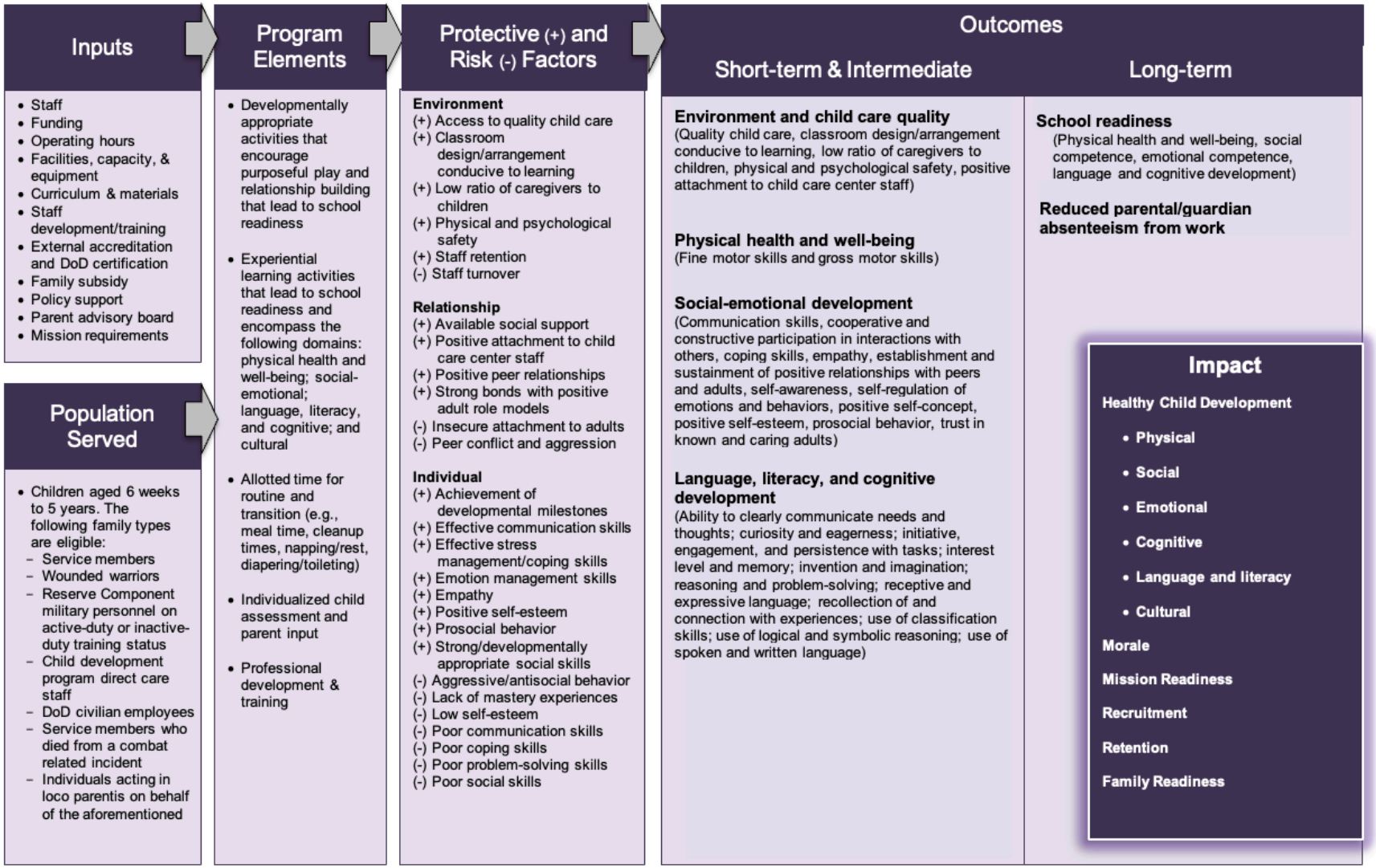
The evaluation was an outcome-focused evaluation, which means the evaluation was intended to assess the degree to which participants achieved the intended outcomes of the program. The following four questions guided the evaluation:

- Does the CDC program improve child development outcomes among program participants?
- Is the CDC program more or less effective across different groups of participants (e.g., children who have direct-care staff with more or less training)?
- Do children attending military CDC programming differ from military children in comparable civilian programs? If so, how do they differ?
- How do children involved in CDC programming differ from normative comparison groups?

The evaluation was based on a logic model (see Figure 1) developed during the planning phase of this evaluation (Davenport et al., 2016).

Figure 1

ECE Project Logic Model



Methods

Evaluation Design

The evaluation design was quasi-experimental and compared military children in CDCs and military children in civilian child care centers. Child development was intended to be assessed over the course of a year at 3-month intervals for a total of five timepoints (i.e., five waves). Parent outcome, child psychological well-being, process, demographic, and moderator variable measures were intended to be collected at the baseline, 6 months, and 1-year data-collection timepoints. To minimize burden, only the child development measures were originally scheduled to occur at all five waves. Table 1 outlines the original data-collection schedule.

To retain as many program participants as possible throughout the course of the evaluation, incentives were offered to families at each timepoint. This is a common practice in longitudinal research and evaluation, so the participants remain interested and invested. Parent participants received $15 for completing the first questionnaire, $15 for the second, $25 for the third, $25 for the fourth, and $35 for the fifth and final questionnaire.

The data-collection timeline was based on CDCs' assessment schedules, which varied by Service, and, in some instances, by CDC. The evaluation's five waves of data collection were aligned with the following Teaching Strategies Gold (TS Gold) assessment periods: Summer 2018/2019 (Cohort 1 only), Fall 2019/2020, Winter 2019/2020, Spring 2019/2020, Summer 2019/2020, and Fall 2020/2021 (Cohort 2 only).

Table 1
Original Data-Collection Schedule

	Wave 1: Baseline	Wave 2: 3 months	Wave 3: 6 months	Wave 4: 9 months	Wave 5: 12 months
Parent Report					
Child Development	X	X	X	X	X
Child Well-being	X		X		X
Family-Work Conflict	X		X		X
Parental Absenteeism from Work	X		X		X
Parent Perceived Stress	X		X		X
Demographic/Process Variables	X		X		X
Staff Report					
Child Development	X	X	X	X	X
Classroom Environment Questionnaire	X		X		X
Director Report					
Leadership Support	X		X		X
Observer Report					
Environment Rating Scale	X		X		

COVID-19 Pandemic

In mid-March 2020, the child care centers involved in this evaluation closed in response to the global COVID-19 pandemic. At the same time, the evaluation team was preparing the requests for the Winter 2019/2020 child assessments and Wave 3 Classroom Environment Questionnaires. When the CDCs closed in March, at the direction of OSD, the Clearinghouse evaluation team ceased communication with the CDCs.

Although some of the CDCs remained open to mission-essential families, others closed entirely. The 12 CDCs involved in this evaluation reopened at different times with different statuses (e.g.,

mission essential only, fully open). When the CDCs did reopen, there were no transitions to new classrooms, new enrollments did not occur, and assessments were put on hold.

In August 2020, the decision was made by OSD, in partnership with the Clearinghouse evaluation team, to end data collection from the CDCs (i.e., child assessments, direct-care staff-reported Classroom Environment Questionnaire, director questionnaires, and the second observation). Data collection from the parents continued through Wave 5 as originally intended. In August 2020, the evaluation team began the process of requesting the reports for the child assessments that were completed by the CDCs prior to the March 2020 closures. Retrieval of those reports concluded in April 2021. Ultimately, director-reported questionnaires, the direct-care staff-reported Classroom Environment Questionnaire, and the observations ended after Wave 1. Collection of the TS Gold direct-care staff child assessments concluded after Wave 3 for Cohort 1 and Wave 2 for Cohort 2. To reduce burden on families, the evaluation team had originally reduced collection of some of the parent measures (e.g., family-work conflict, perceived stress) to three of the five waves of data collection (i.e., Wave 1, Wave 3, and Wave 5). After the onset of the COVID-19 pandemic, capturing these specific measures at all subsequent waves was important. Therefore, the evaluation team added those measures to the Wave 4 questionnaire.

During the global COVID-19 pandemic, many parents began working from home while also caring for their children. Other parents left the workforce to care for their children. This may have affected parent responses to questionnaires in multiple ways that the evaluation team was unable to measure. First, parents may have spent a lot more time with their children than they had previously spent, which may have resulted in a different understanding of their children's abilities. Second, children's behaviors may have been impacted by the disruption caused by the pandemic. Third, parents may have been unusually stressed; thus, they may have perceived their child's behavior differently than they typically would have.

The actual data-collection schedule is presented in Table 2. It can be contrasted with the originally planned data-collection schedule, which is presented on page 9.

Table 2
Actual Data-Collection Schedule

	Wave 1: Baseline	Wave 2: 3 months	Wave 3: 6 months	Wave 4: 9 months	Wave 5: 12 months
Parent Report					
Child Development	X	X	X	X	X
Child well-being	X		X	+	X
Family-Work Conflict	X		X	+	X
Parental Absenteeism from Work	X		X	+	X
Parent Perceived Stress	X		X	+	X
Demographic/Process Variables	X		X	+	X
Staff Report					
Child Development	X	X	X / –	–	–
Classroom Environment Questionnaire	X		–		–
Director Report					
Leadership Support	X		–		–
Observer Report					
Environment Rating Scale	X		–		

Note. X indicates that the data collection occurred as planned. + indicates that, due to the COVID-19 pandemic, the measures were added to the data-collection timepoint. – indicates that, due to the COVID-19 pandemic, data were not collected as planned. X / – indicates that data were collected as planned for Cohort 1, but data were not collected for Cohort 2.

Measures

Parent Completed Measures

A parent or guardian completed the Ages and Stages Questionnaire-3 (ASQ-3; Squires & Bricker, 2009), the Ages and Stages Questionnaire: Social Emotional-2 (ASQ:SE-2; Squires et al., 2010), the Early Development Instrument (EDI; Janus & Offord, 2007), the Family Work Conflict scale

(FWC; Netemeyer et al., 1996), the Strengths and Difficulties Questionnaire (SDQ; Goodman et al., 1998), the Perceived Stress Scale (PSS-10; Cohen & Williamson, 1988), and a single-item parental absenteeism from work measure. In addition, parents completed demographic information and process items. Reliability and validity information for the measures are available in Appendix A.

Demographic and process information. Parents provided basic demographic information including their own age, education, employment status, gender, spousal employment status (if applicable), and relationship status. Parents were also asked to provide specific military-related demographics including Service member rank, military service history (e.g., deployment, relocation), and current military status. In addition, parents indicated how many hours per week their child attended the child care center, how long their child had attended the child care center, and child-specific demographic information (e.g., age, gender, race, special needs status).

ASQ-3. The ASQ-3 is a 30-item screening instrument used to assess whether children aged 1 to 66 months display typical child development (Squires & Bricker, 2009). The ASQ-3 consists of 21 age-appropriate questionnaires (e.g., 22 Month Questionnaire for children 21 months 0 days through 22 months 30 days). The instrument consists of five subscales (i.e., communication, gross motor, fine motor, problem-solving, and personal-social), and each subscale contains six items. The responses for each item are “yes,” “sometimes,” or “not yet.” After scale scores were calculated, for each subscale, children were placed in one of three categories: typical development, monitor, refer.

ASQ:SE-2. The ASQ:SE-2 is a screening instrument used to assess whether children aged 1 to 72 months display typical child development in the social-emotional domain (Squires et al., 2010). The ASQ:SE-2 complements the ASQ-3 but can also be used independently. The ASQ:SE-2 consists of nine age-appropriate questionnaires (e.g., 24 Month Questionnaire for children 21 months 0 days through 26 months 30 days). The instrument has 21 to 32 items depending on the version of the questionnaire. Questionnaires for older children have more items. Each item on the questionnaire asks parents to rate the frequency with which specific behaviors or skills occur (i.e., Often or Always, Sometimes, Rarely or Never). Once the scale score is calculated, children were placed into one of three categories: no or low risk, monitor, refer.

EDI. The EDI is a 103-item measure of child development for children between 4 and 5 years of age (Janus & Offord, 2007). The EDI measures five domains of child development: physical health and well-being, social competence, emotional maturity, language and cognitive development, and communication skills and general knowledge. The EDI was adapted for use in the ECE Project. This adaptation is 31 items and only examines school readiness in the cognitive, communication, and physical domains. In addition, this measure was originally developed as a teacher report of school readiness. In the ECE Project, the evaluation team used the EDI as a parent report measure due to the utility of the items and the lack of parent report measures for school readiness. Subscale scores range from 0-10, and higher scores reflect greater development.

SDQ. The SDQ, developed by Goodman and colleagues (1998), is a 25-item instrument used to assess psychological adjustment of children 2 to 17 years old. It contains five subscales:

emotional problems, conduct problems, hyperactivity, peer problems, and prosocial behavior. Parents rate on a 3-point scale the extent to which each of the 25 items applies to the child of interest. Scores range from 0 to 40 (i.e., total difficulties), 20 (i.e., internalizing and externalizing scales), or 10 (i.e., individual subscale scores). For all scales, except the prosocial behavior subscale, higher scores reflect more problematic functioning. For the prosocial behavior subscale, higher scores reflect better functioning. SDQ scores were also transformed into categories based on the normed data. For the total difficulties score and the individual subscales, children were placed into one of four categories: close to average, slightly raised/lowered (i.e., lowered for the prosocial subscale, raised for all others), high/low, and very high/low.

FWC. The FWC is a 5-item scale that assesses whether the family role interferes with work-related responsibilities (Netemeyer et al., 1996). The response options are on a 7-point scale that ranges from 1 (strongly disagree) to 7 (strongly agree). Responses are averaged to create a continuous scale score that ranges from 1-7. Higher scores reflect greater family-work conflict.

PSS-10. The PSS-10 is a 10-item measure that assesses the degree to which an individual perceives aspects of his or her life as uncontrollable, unpredictable, and overwhelming (Cohen & Williamson, 1988). This is a parent self-report measure. The measure asks the individual to reflect upon the past month. Each question is rated on a 5-point scale from 0 (never) to 4 (very often). Items are summed to create a total perceived stress score that ranges from 0 to 40, and higher scores reflect greater stress. In addition to the total score, two subscales were also calculated, and these created a score for perceived helplessness and perceived self-efficacy. Scores range from 0 to 24 and 0 to 16, and higher scores reflect greater helplessness and less self-efficacy, respectively.

Parental Absenteeism. The single-item parental absenteeism measure is based on an item in the Defense Manpower Data Center Status of Forces Survey of Active Duty Members (SOFS-A) survey (C. Stevens, personal communication, March 24, 2016). Participants were asked, “In the past 6 months, how many days of work have you missed because of a lack of child care?”

CDC Staff Completed Measures

A direct-care staff member completed the TS Gold child assessment. A CDC staff member with knowledge of the classroom (e.g., direct-care staff, training and curriculum specialist) completed measures related to program processes.

Classroom Environment Questionnaire (CEQ). CDC staff provided information on the classroom direct-care staff, including their education/credentials, duration of employment at the current installation, duration of employment in their current role, and duration of employment in the child care field. CDC staff also provided information on the number of staff and children in the classroom. Three items related to leadership support were also assessed. Direct-care staff reported, on a 5-point scale (i.e., strongly disagree, disagree, neither agree nor disagree, agree, strongly), their perception of the following items: 1) The director is effective in addressing the

needs of the direct-care staff; 2) The director is effective in addressing the needs of enrolled children and families; and 3) Command at this installation is supportive of the CDC.

Director-reported Leadership Support. The CDC directors provided their perception of command/leadership support for the CDC program. Directors reported, on a 5-point scale (i.e., strongly disagree, disagree, neither agree nor disagree, agree, strongly), their perception of the following items: 1) Command at this installation is supportive of the CDC and 2) Command at this installation is supportive of my role as director.

TS GOLD. The TS GOLD is a 51-item authentic, observation-based measure that is designed to assess children from birth through kindergarten (Heroman et al., 2010). It is the child assessment regularly used by CDC staff to assess the progress of the children in the program. The measure focuses on indicators of school readiness (i.e., cognitive, language, literacy, mathematics, physical, and social-emotional development). CDC staff rated children on each domain using a 10-point scale from 0 to 9, and they had item-specific response options to guide their rating. A widely held expectation band is available for each item and varies by item within a subscale. These widely held expectations are based on normed data and indicate whether a child is progressing towards expectations, meeting expectations, or exceeding expectations. A final score of progressing towards expectations, meeting expectations, or exceeding expectations is calculated for each subscale (i.e., social-emotional, physical, language, cognitive, literacy, and math).

The Services use both the online and the paper versions of this assessment. The two versions, while conceptually similar, are not identical. Because the online version included items that were not in the paper version and since published reliability and validity data were based on the paper version, the evaluation team matched the online version to the paper version. Furthermore, some of the items had different wording, despite having conceptually similar meaning. As such, the team conducted analyses of invariance to determine if the difference in wording between versions made a difference in how the items were scored. The analyses found that for the language and literacy subscales, the wording differences mattered. Therefore, in the outcome analyses, the evaluation team controlled for the effects of the different versions of the measure.

Observer Completed Measures

ERS. The ERS includes two measures: the Infant and Toddler Environment Rating Scale, 3rd Edition (ITERS; Harms et al., 2017) and the Early Childhood Environment Rating Scale, 3rd Edition (ECERS; Harms et al., 2015). The ERS assesses structural aspects of the child care environment. The ERS can be used as an ongoing self-assessment and as a program-evaluation tool.

The ITERS is designed to assess classrooms that serve children who are birth to 3 years old (Harms et al., 2017) and contains 33 items organized into the following six subscales: Space and Furnishings, Personal Care Routines, Language and Books, Activities, Interaction, and Program Structure. Subscale scores range from 1 to 7, and higher scores reflect better quality.

The ECERS is designed to assess classrooms of children who are 3 to 5 years old (Harms et al., 2015). The measure contains 35 items organized into the following 6 subscales: Space and Furnishings, Personal Care Routines, Language and Literacy, Learning Activities, Interaction, and Program Structure. Subscale scores range from 1 to 7, and higher scores reflect better quality.

Observer Selection and Training. An independent observer (i.e., a Penn State staff member) completed the Environment Rating Scale (ERS). Using an external observer minimizes potential bias and staff burden. Observers were sought for 10 geographic locations. Advertisements were placed on Indeed.com in the 10 geographic areas where the CDCs and the civilian centers were located. In addition, the advertisement was shared via social media and military spouse networks. When possible, observers were hired in the geographic location of the centers. However, qualified candidates were not found for all locations, and, therefore, observers traveled to those locations to conduct observations. Observers had backgrounds in early childhood education, child development, psychology, or data collection. Eight individuals were hired for the first observation. An additional three people were hired for the second observation; however, the second observation did not occur due to the COVID-19 pandemic. Seven of the eleven observers who were hired were current or former military spouses.

Observers who conducted observations completed training at Penn State. At this training, individuals learned the indicators, items, and scoring system; they had the opportunity to ask questions; they scored video examples of classrooms and interactions; they participated in practice observations in child care center classrooms in State College, Pennsylvania.

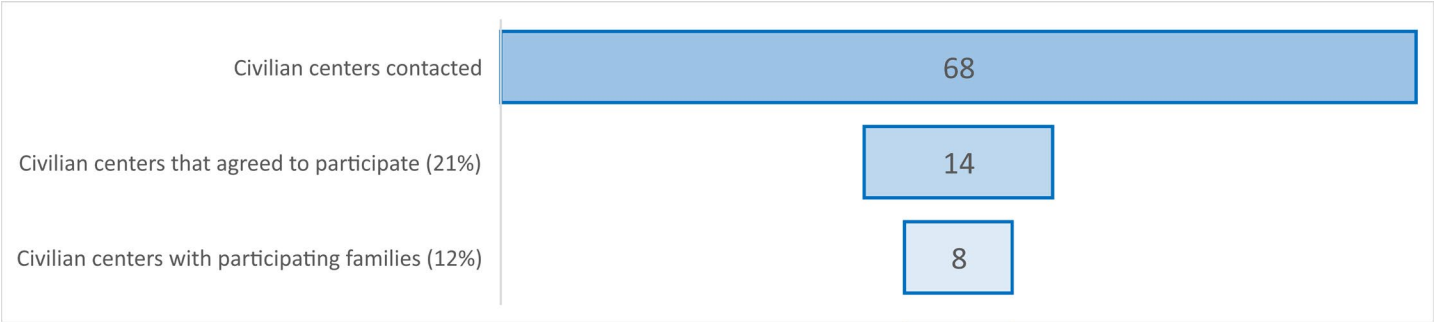
Recruitment and Registration

Recruitment

Headquarters points of contact (POCs) from the Army, Navy, and Marine Corps[1] each selected four CDCs to participate in the evaluation. These 12 CDCs were located at 10 different installations in 9 geographic locations. POCs were provided with a list of several factors to consider when selecting CDCs to participate. These elements included program size, the number of families in the area who use civilian centers, whether the installation is demographically representative of the military, and agreeableness of the CDC.

In order to recruit military families who were utilizing civilian centers, the Clearinghouse contacted 68 civilian child care centers. The criteria for contacting the civilian centers were as follows: (1) centers were identified by the Army, Navy, and Marine Corps POCs as participating in the fee-assistance program; (2) centers were within 35 miles (45 miles in 2 areas with a large number of military installations) of a participating CDC; and (3) the center had five or more military families who were participating in the DoD fee-assistance program. To increase the number of families from civilian centers who could participate in the evaluation, this last criterion was relaxed as more centers were contacted. Figure 2 presents the number of civilian child care centers that were contacted, the number that agreed to participate, and the number of centers with participating families.

Figure 2
Number of Civilian Child Care Centers



Recruitment of families began in July 2019 and ended in December 2019. The CDCs and the civilian centers were instrumental in helping recruit families to participate in the evaluation. The evaluation team developed flyers, posters, FAQs, social media posts, an email message, and a newsletter blurb for CDC and civilian center staff to use to help recruit families for the evaluation.

[1] The Air Force choose not to participate in the evaluation because they were implementing a new curriculum at the time and were concerned about staff burden.

In addition, a project website was developed to provide information about the evaluation. The website also contained a link to the registration form for the evaluation. All recruitment material included information that explained how to contact the evaluation team with questions. Due to the low recruitment numbers for the civilian centers, Clearinghouse evaluation staff traveled to four civilian centers in an attempt to increase the recruitment numbers. Some of the civilian centers chose not to allow in-person recruitment by Clearinghouse staff.

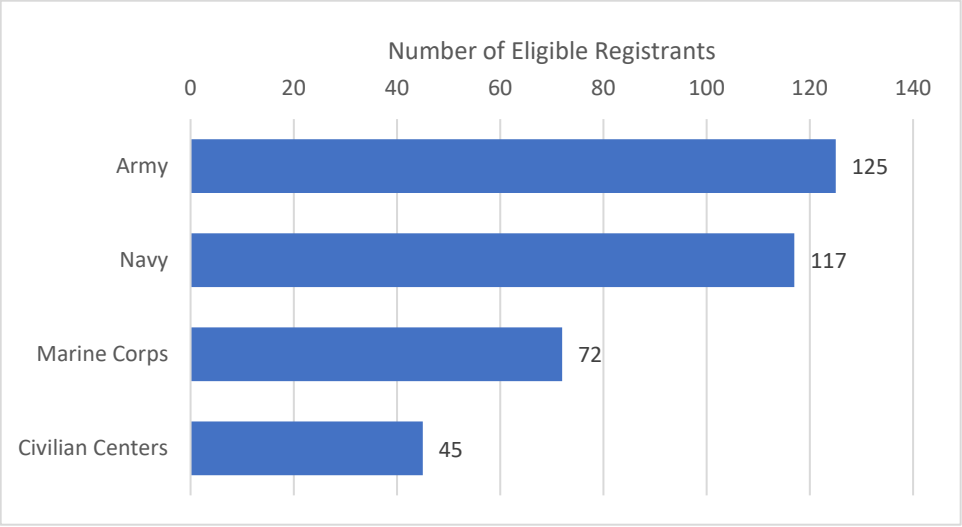
Recruitment materials were also provided for center staff. This included a memo and an FAQ to help explain the evaluation and proactively answer questions about the evaluation. A section for staff also appeared on the project website, which included the FAQ, instructions (i.e., a video and a PDF) for how to print the reports that were requested for the evaluation, and information on how to contact the evaluation team with questions.

Per guidance from Penn State's Institutional Review Board, since this was program evaluation and not research, a consent form was not necessary. However, all parents, direct-care staff, and directors were provided with an information sheet, which contained the information that is typically provided on a consent form. This information sheet was provided with the first questionnaire and was available on the website.

Registration

A total of 483 registrations were recorded in the registration system. There were 359 eligible registrations from Army, Navy, Marine Corps, and civilian child care centers. The number of eligible registrations per Service and for civilian centers is shown in Figure 3.

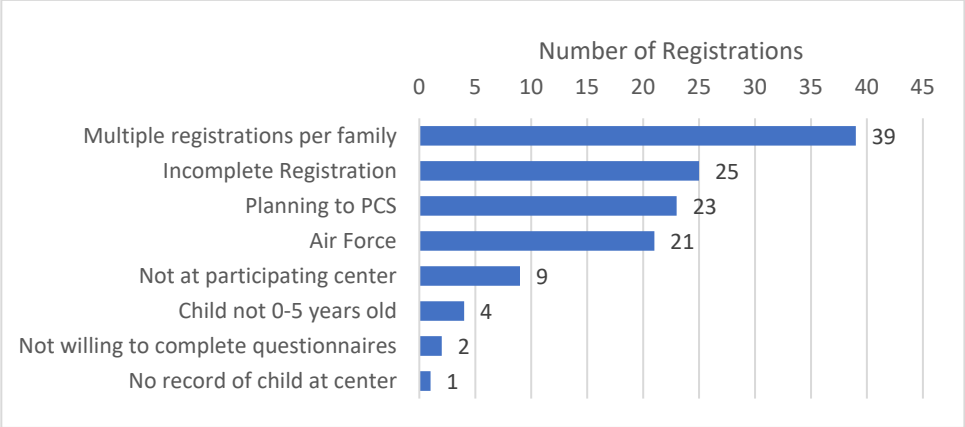
Figure 3
Child Care Center Service Affiliation for Eligible Registrants (n=359)



Twenty-five registrations were incomplete, and 99 were ineligible. Eligibility criteria were outlined during the registration process; families had to affirm that they met all eligibility criteria. Some ineligible registrants were identified at registration and were notified automatically through the registration system that they were not eligible. Some ineligible registrants were identified after the registration process and were notified via email that they were not eligible. The reasons for ineligibility are listed in Figure 4.

Figure 4

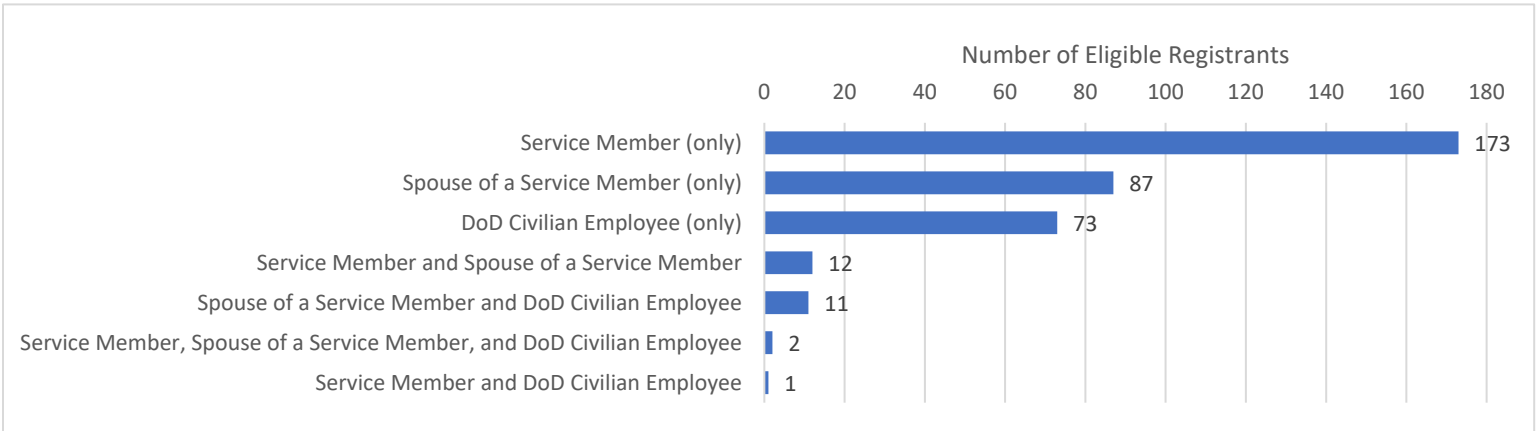
Number of Incomplete and Ineligible Registrations and Reason for Ineligibility (n=124)



Registrants could be active duty Service members, spouses of active duty Service members, DoD civilian employees, or a combination thereof. The largest number of registrants were Service members. Figure 5 shows the categories that were selected by the registrants. Three registrants indicated that they were both an active duty Service member and a DoD civilian employee; they are listed separately in Figure 5.

Figure 5

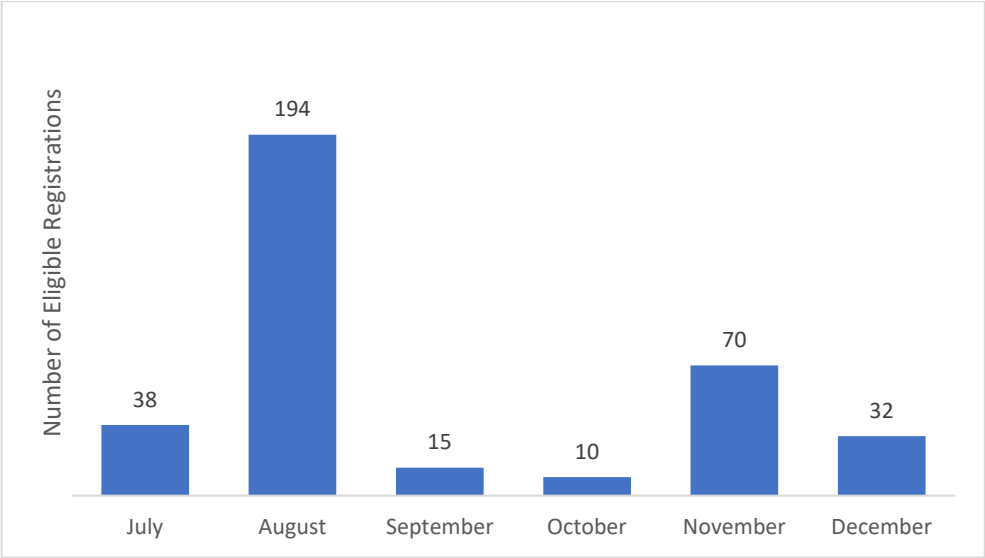
DoD Affiliation of Eligible Registrants (n=359)



Registration for the project opened in July 2019 and closed in December 2019. The first recruitment push occurred in July/August 2019. The second push occurred in November 2019. The timing of the recruitment pushes is reflected in the large number of registrations in August and November as shown in Figure 6.

Figure 6

Number of Families Who Registered Each Month (n=359)



A total of 314 families who were eligible at registration attended CDCs. The remaining 45 families attended civilian centers. Table 3 displays the number of children who were involved in the evaluation and the number of classrooms in which these children were enrolled for each Service and for the civilian centers.

Table 3

Number of Children in Participating Families per Service and Number of Associated Classrooms

Service	Children	Classrooms
Army	125	50
Navy	117	49
Marine Corps	72	43
Civilian	45	28
Total	359	170

Table 4 illustrates the number of registered children in eligible families and associated classrooms for the CDCs. As shown in the table, 314 children were enrolled in 142 classrooms. On average, there were 26 children per CDC who were involved in the evaluation, and those children were in an average of 12 classrooms per CDC. However, the range was large, with between 8 and 74 children per CDC and between 4 and 22 associated classrooms per CDC. On average, 11% of the children in each CDC were included in the evaluation, and these children were, on average, in 70% of the classrooms in each CDC. The range was large. Between 4% and 25% of all children at each CDC were involved in the evaluation, and between 35% and 100% of all classrooms at each CDC were included in the evaluation.

Table 4

Total and Average Number of Children and Classrooms in CDCs

12 CDCs	Children	Classrooms
Total number	314	142
Average (#) Per CDC	26	12
Range (#)	8 to 74	4 to 22
Average (%) Per CDC	11%	70%
Range (%)	4% to 25%	35% to 100%

Table 5 shows the number of children in families who were eligible at registration and the number of associated classrooms for civilian centers, the average number of children and classrooms per center, and the highest and lowest number of children and classrooms per center. The percentage of children at each civilian center who were included in the evaluation is not included in Table 5, as it was for Table 4, because these data were not available.

Table 5

Total and Average Number of Children and Classrooms in Civilian Centers

8 Civilian Centers	Children	Classrooms
Total number	45	28
Average (#) Per Center	6	4
Range (#) Per Center	2 to 11	2 to 5

Attrition, withdrawal, and disenrollment

Between registration and Wave 1, an additional 29 families became ineligible or withdrew from the evaluation (see Figure 7). Because data were only collected from the parents at the civilian centers (i.e., not the direct-care staff), if parents at civilian centers did not complete the Wave 1 questionnaire, they were removed from the evaluation. In addition, due to a technical issue, some

parents received the permission form at the beginning of the Wave 1 questionnaire instead of during registration. For parents at CDCs who received the permission form at the beginning of Wave 1 and did not complete the permission form or the Wave 1 questionnaire, direct-care staff-reported data were not collected, and those families were considered ineligible. For families at CDCs who completed the permission form at registration and did not complete the Wave 1 questionnaire, they did not receive subsequent parent questionnaires; however, data were still collected from the direct-care staff. There were 328 families who were considered eligible participants at the completion of Wave 1 of the evaluation (see Figure 8).

Figure 7

Registrants Who Were Eligible at Registration but Became Ineligible or Withdrew from the Evaluation at or Before Wave 1

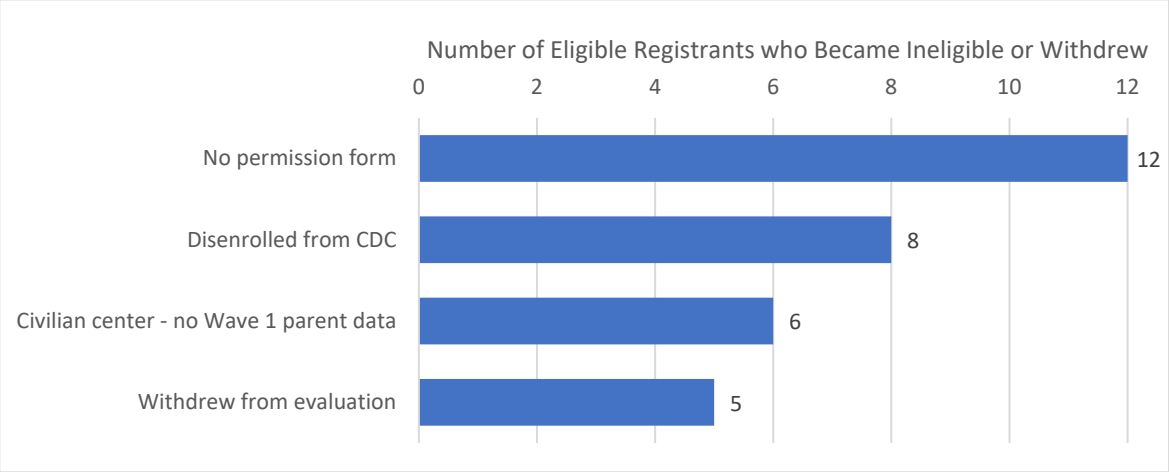
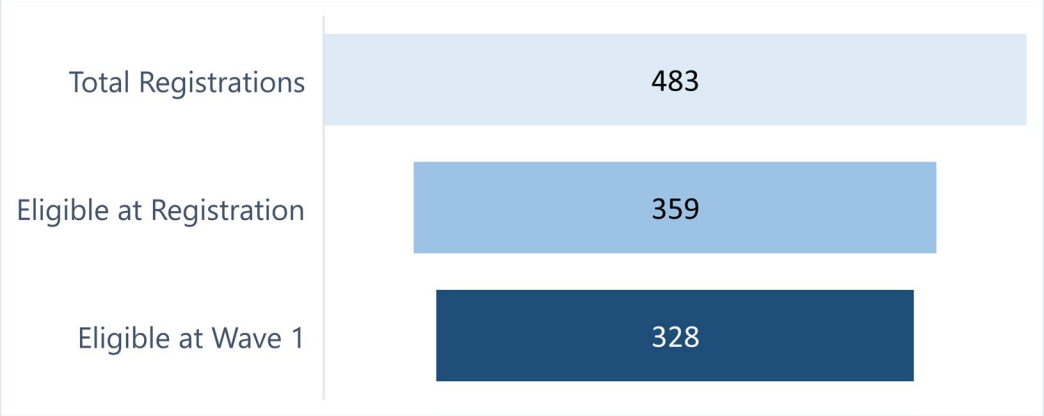


Figure 8

Evaluation Registration and Eligibility



Participants left the evaluation for multiple reasons and at different times. In addition, for parents at CDCs, those who did not complete the Wave 1 questionnaires did not receive subsequent questionnaires. However, because their children had direct-care staff-reported data, they were still enrolled in the evaluation. Conversely, for parents at civilian centers, because there was no direct-care staff-reported data from the civilian sites, those who did not complete the Wave 1 questionnaires were removed from the evaluation. As such, the response rate and eligibility status data are complicated. Table 6 outlines the number of families who were eligible and who received questionnaires at each wave. The number of families who were eligible after Wave 1 remains constant regardless of whether families disenrolled from the CDC or did not complete subsequent questionnaires; all families who have data for at least one wave from at least one reporter (i.e., parent or direct-care staff) remain eligible and part of the evaluation.

Table 6
Family Evaluation Eligibility, Parent Questionnaire Distribution, and Parent Questionnaire Responses at Waves 1, 2, 3, 4 and 5

	# of families
Total number of families eligible at registrations	359
Disenrolled from CDC between registration and Wave 1	1
Number of families who received the Wave 1 questionnaire	358
Number of parent responses to Wave 1 questionnaire	266
Number of families who became ineligible at Wave 1	30
No permission form received	[12]
Disenrolled from CDC - no data were collected	[7]
No Wave 1 parent data received - family at civilian center	[6]
Withdrew from evaluation (includes 2 parents who completed the Wave 1 questionnaire)	[5]
Number of eligible families after Wave 1 (data are available from at least one reporter for at least one timepoint)	328
Eligible families who did not receive the parent questionnaire after Wave 1	66
Did not respond to the Wave 1 questionnaire	[60]
Disenrolled from the CDC after Wave 1	[6]
Number of families who received the Wave 2 questionnaire	262
Number of parent responses to Wave 2 questionnaire	227
Eligible families who did not receive the Wave 3 or later questionnaire	14
Disenrolled form CDC after Wave 2	[14]
Number of families who received the Wave 3 questionnaire	248
Number of parent responses to Wave 3 questionnaire	207
Eligible families who did not receive the Wave 4 or later questionnaire	1
Disenrolled from CDC after Wave 3 but before COVID	[1]
Number of families who received the Wave 4 questionnaire	247
Number of parent responses to Wave 4 questionnaire	197
Number of families who received the Wave 5 questionnaire	247
Number of parent responses to Wave 5 questionnaire	188

Some parents only completed the first question on the questionnaire and, therefore, were not included in the final data set. The number of parents for whom this was the case is as follows: Wave 1 = 1; Wave 2 = 0; Wave 3 = 4; Wave 4 = 1; Wave 5 = 3.

Data Collection

Parents

The timing of parent questionnaire distribution was based on the CDCs' assessment schedules. For parents at CDCs, the parent questionnaires were distributed 1 week after the TS Gold direct-care staff assessment deadline. For parents in the civilian centers, the questionnaires were distributed at the same time as they were distributed to the parents in the CDCs in the same geographic location. The original intent was to send one reminder to parents 2 weeks after the initial distribution of the questionnaire. However, the Fall 2019 (Cohort 1 Wave 2 and Cohort 2 Wave 1) reminder was received during the December holidays. Therefore, in consultation with OSD, a second reminder was added; it was to be distributed 4 weeks after the initial questionnaire distribution. Subsequently, the Winter 2020 (Cohort 1 Wave 3 and Cohort 2 Wave 2) assessment occurred at the beginning of the COVID 19 pandemic. At that time, a third reminder was added; it was distributed 6 weeks after the initial questionnaire distribution. This expanded reminder distribution remained in effect for the subsequent the data-collection timepoints. Table 7 displays the number and percent of participants who received reminder emails.

Table 7

Number of Questionnaires and Reminders Sent (% of Respondents Who Received the Reminder) for the Final Sample

	Wave 1 # (%)	Wave 2 # (%)	Wave 3 # (%)	Wave 4 # (%)	Wave 5 # (%)
Questionnaires sent	328	262	248	247	247
Reminder 1 sent	140 (43%)[a]	105 (40%)[a]	112 (45%)[a]	116 (47%)[a]	130 (53%)[a]
Reminder 2 sent	31 (9%)[b]	61 (23%)[a]	70 (28%)[a]	71 (29%)[a]	86 (35%)[a]
Reminder 3 sent	-	14 (5%)[b]	48 (19%)[a]	62 (25%)[a]	69 (28%)[a]

Note. [a] = reminder received by Cohort 1 and Cohort 2; [b] = reminder received by only Cohort 2

Overall, parents responded in a timely fashion to the questionnaires. Many parents responded the same day that they received the questionnaire or the reminder. The evaluation team's target was for parents to complete the questionnaire within 4 weeks of the original distribution. Of the parents who responded to the questionnaire, an overwhelming majority responded within 4 weeks: Wave 1 = 98%, Wave 2 = 94%, Wave 3 = 94%, Wave 4 = 94%, Wave 5 = 93%. Table 8 shows the number and percent of respondents who responded within each original distribution or reminder period.

Table 8

Response Time for Participants in the Final Sample Who Responded to the Questionnaire

	Wave 1 # (%)	Wave 2 # (%)	Wave 3 # (%)	Wave 4 # (%)	Wave 5 # (%)
Responded after initial distribution	190 (72%)	162 (71%)	137 (66%)	139 (71%)	117 (62%)
Responded after reminder 1	63 (24%)	43 (19%)	42 (20%)	38 (19%)	49 (26%)
Responded after reminder 2	10 (4%)	17 (7%)	24 (12%)	10 (5%)	15 (8%)
Responded after reminder 3	1 (0%)	5 (2%)	4 (2%)	10 (5%)	7 (4%)

Wave 1 questionnaires were sent to all families who were eligible at the time of the original Wave 1 distribution. Questionnaires for subsequent waves were sent only to those parents who completed the Wave 1 questionnaire and who remained eligible for the project. Until March 2020, families who disenrolled from a participating child care center became ineligible to complete the evaluation. Their data from timepoints prior to their disenrollment remained included in the evaluation, but they were not sent additional questionnaires. However, beginning in March 2020, due to the disruption caused by the COVID 19 pandemic, parents were sent questionnaires regardless of their current child care center enrollment status.

Table 9 displays the response rates for the parent questionnaires. Note, for those parents who did not complete the Wave 1 questionnaire, although they remained in the evaluation[2], they were not sent further questionnaires, and, therefore, they are not included in the response rate calculation in Table 9. Furthermore, some families were eligible to participate at registration and received questionnaires but were removed from the final sample. These families (n=31) either disenrolled from the child care center before data could be collected, asked to withdraw from the evaluation, did not complete the Wave 1 questionnaire and did not complete the permission form allowing the evaluation team to collect the direct-care staff assessment data, or were enrolled in a civilian center and did not complete the Wave 1 parent questionnaire. Please note, not everyone who responded to the questionnaire answered all of the questions. Table 10 displays the percent of families who were eligible at registration and who responded to the questionnaires.

Table 9

Response Rates for Sent Parent Questionnaires for the Final Sample (n=328)

	Number of Questionnaires Sent	Number of Responses	Response Rate
Wave 1	328	264	80%
Wave 2	262	227	87%
Wave 3	248	207	83%
Wave 4	247	197	80%
Wave 5	247	188	76%

[2] These families remained in the evaluation because direct-care staff-reported data were available.

Table 10
Questionnaire Response Rates for all Participants Eligible at Registration (n=359)

	Response Rate
Wave 1	74%
Wave 2	63%
Wave 3	58%
Wave 4	55%
Wave 5	52%

Directors

In an effort to better understand how the directors' perceptions of leadership support might play a role in child development or school readiness, directors were sent a two-item online survey. They were asked to rate, on a scale of 1 (i.e., Strongly Disagree) to 5 (i.e., Strongly Agree), the following questions: 1) Command at this installation is supportive of the CDC. 2) Command at this installation is supportive of my role as director. This survey was sent to the 12 directors at the participating CDCs. Seven complete responses were received. One director opened the survey but did not answer the questions. Four directors did not open the survey. Reminders were sent to the seven directors who had not interacted with the survey after 2 weeks. Director responses were received between 0 and 43 days after distribution. Table 11 shows the response rate, the minimum and maximum number of days that passed between when the survey was sent and when it was completed, and the average number of days it took directors to complete the survey.

Table 11
Director Questionnaire Response Rate

	Director Responses
Response Rate	58%
Minimum Days to Complete	0 days
Maximum Days to Complete	43 days
Average Days to Complete	13 days

Direct-care Staff

CEQ

At the end of the first assessment period (i.e., Summer 2018/2019 for Cohort 1 and Fall 2019/2020 for Cohort 2), 139 classroom direct-care staff received the CEQ (i.e., one per classroom). A total of 122 completed CEQs were received by the evaluation team. These responses were received between 12 and 148 days after they were mailed. Between 4 and 22 classrooms at each center received the questionnaires. Response rates were generally high but varied by CDC. For all but

one CDC, response rates ranged from 80-100%. For one CDC, there was a misunderstanding, and only 68% of the CEQs were completed and returned. Table 12 outlines the response rates for the CEQs.

Table 12
CEQ Response Rates

	CEQ Responses
Response Rate - Overall	
% Completed and Returned	88%
Minimum Days to Receive Completed CEQ	12 days
Maximum Days to Receive Completed CEQ	148 days
Average Days to Receive Completed CEQ	33 days
Response Rate – Per CDC	
Minimum % Completed at Each CDC	68%
Maximum % Completed at Each CDC	100%
Average % Completed at Each CDC	89%

TS Gold

CDC staff conduct child assessments four times per year using the TS Gold. Assessments are conducted either online (for Army and Navy) or using a paper portfolio (for Marine Corps). Although the intention was to collect the data from these assessments five times for each child, because of the COVID 19 pandemic, data were only collected for three timepoints for Cohort 1 and twice for Cohort 2. All direct-care staff-reported child assessments that were collected occurred before the COVID 19 shutdown that occurred in March 2020.

For the online assessments, direct-care staff were asked to print and send reports for assessments they conducted throughout the regular course of their duties. Direct-care staff were provided with the names of the children, the reports that were being requested, and instructions for printing the reports. For the paper portfolio, members of the evaluation team photocopied the portfolios at the beginning of the evaluation. The photocopying of assessment portfolios by members of the evaluation team was also to occur at the end of the evaluation. However, due to the COVID 19 pandemic, this was not possible. In collaboration with the Marine Corps headquarters POC, an alternative solution was established. The evaluation team sent memory cards to the Marine Corps CDCs. CDC staff then took photographs of each page of the portfolio and sent the memory cards back to Penn State.

Due to the COVID 19 pandemic, many children disenrolled from the CDCs between March 2020 and the end of the evaluation period. The paper assessment portfolios were given to the parents when the children disenrolled from the center. Therefore, these portfolios were not available to be

photographed. Similarly, for the online assessments, when the children disenrolled from the centers, the online profiles were archived. Most of the archived profiles were able to be retrieved; however, some were not.

Table 13 outlines the reasons for and number of instances in which reports were not received. Of the 289 CDC families participating in the evaluation, 40 children were missing TS Gold reports at Wave 1, 74 were missing reports at Wave 2, and 159 were missing reports at Wave 3. Several reasons were provided for why reports were not available. 1) Children who were newly enrolled were not assessed because an accurate assessment could not be conducted. 2) Children disenrolled and, therefore, no assessment was conducted. 3) When children disenrolled, some parents were given the paper version of the assessment portfolio and data from previous assessments were not retained by the CDC. 4) When children disenrolled, their profiles were archived; for some children, the reports from previous assessments were not retrievable. 5) Some reports were not available or received by the Clearinghouse, for unspecified reasons. 6) Some assessments were not conducted, for unspecified reasons. 7) Some reports were received for the requested checkpoint; however, there were no data in the report and no indication that the child had disenrolled prior to the assessment. 8) In one instance, the data were unclear due to unclear data notations.

Table 13
Reasons for and Number of Reports not Received

	Wave 1	Wave 2	Wave 3
Child was newly enrolled	6	0	0
Child disenrolled before assessment period	5	21	21
Portfolio given to parent	7	31	26
Profile archived and not retrievable	2	3	1
Not available/not received (unspecified)	9	8	7
Not assessed (unspecified)	0	4	1
Report received but had no data for requested checkpoint	10	7	4
Unclear checkpoint notations	1	0	0
Cohort 2 Wave 3 – no data due to COVID 19	-	-	99
Total # (%) of children in the final sample missing all TS Gold data for the Wave	40 (14%)	74 (26%)	159 (55%)

For reports that were received and had data, some had items that were missing. Because analyses are conducted at the aggregated subscale level, missing data on received reports can affect whether the data can be used in the analysis. The number of children with missing data on the TS Gold reports are listed in Table 14 and are separated by percent of missing items and wave of data collection. At Wave 1, 16% of the reports had some data but were not complete; at Wave 2, 10% of the reports had some data but were not complete; at Wave 3, 6% of the reports has some data but were not complete.

Table 14

Completion Status of TS Gold Reports

	Wave 1	Wave 2	Wave 3
# of participating families enrolled in CDCs in pre-COVID waves	289	289	190
Reports received at 100% completion	202 (70%)	187 (65%)	118 (62%)
Reports received at 76-99% completion	23 (8%)	14 (5%)	5 (3%)
Reports received at 51-75% completion	11 (4%)	9 (3%)	4 (2%)
Reports received at 26-50% completion	9 (3%)	1 (<1%)	0 (0%)
Reports received at 1-25% completion	4 (1%)	4 (1%)	3 (2%)
Reports not received or 0% complete	40 (14%)	74 (26%)	60 (32%)

Note: Due to rounding, totals may not equal 100%.

In order to examine the potential utility of data imputation for the TS Gold, the evaluation team explored missing data at the subscale level. For all reports that had at least one response, the team reviewed the number of responses missing from each subscale for each participant's report. Table 15 presents the number of reports that are missing one to two items and more than two items on each subscale. These data do not include reports that were received with no responses.

Table 15

Number of Reports with Missing Data on Each Subscale for Each Wave

	Wave 1		Wave 2		Wave 3	
Subscale	Missing 1-2 items	Missing more than 2 items	Missing 1-2 items	Missing more than 2 items	Missing 1-2 items	Missing more than 2 items
Social-Emotional	14	9	6	4	1	1
Physical	6	6	5	2	2	2
Language	11	9	8	3	0	1
Cognitive	6	13	3	8	0	1
Literacy	5	7	3	3	2	3
Mathematics	7	5	1	1	2	1

Based on guidance from Lambert (2020), the evaluation team imputed values for missing items in subscales in which one or two items were missing. However, the actual imputation method that was used in Lambert was unclear. In addition, email communication with Lambert and one of his colleagues at Teaching Strategies was unclear. So, the evaluation team followed the theory behind Lambert's imputation method, which stressed the ordinal (i.e., uneven intervals between response options), tripartite (i.e., progressing towards expectations, meets expectations, and exceeds expectations) nature of the responses and the clustered nature of the responses, in general, for individual participants (i.e., they tend to score at either below, at, or above

expectations for most or all items in a subscale). As such, the method used to impute the data is as follows:

1. For each participant, reviewed the non-missing data in each subscale with one or two missing items.
2. Identified whether the majority of the items in the subscale were below, at, or above expectations.
3. Placed each missing item within that same category for that particular missing item.
    a. If the majority of items were *below expectations*, the missing item was imputed as the highest option that was in the *below expectations* category for that item.
    b. If the majority of items were *at expectations*, the missing item was imputed as the middle option that was in the *at expectations* category for that item.
        i. If there were two response options in the category, the item was imputed as the lower of the two options.
        ii. If there were four response options in the category, resulting in two responses being in the middle, the item was imputed as the higher of the two options.
    c. If the majority of items were *above expectations*, the missing item was imputed as the lowest option that was in the *above expectations* category for that item.

A total of 120 items were imputed across the three waves of data. The number of imputed items, the number of participants with imputed items, and the number of computed subscales scores that included imputed items are shown in Table 16.

Table 16

Number of Imputed Items, Number of Participants with Data Imputations, and Number of Computed Subscales with Data Imputations

	Wave 1	Wave 2	Wave 3
# of items imputed	71	39	10
# of participants with data imputations	34	20	6
# of computed subscales with data imputations	49	26	7

For participants with at least some data in the report, a total of 39 participants were missing more than two items on at least one subscale for at least one wave. Participants had between one and six subscales excluded due to missing data. Table 17 provides the total number of participants at each wave with excluded subscales and the number of participants missing each subscale.

Table 17

Number of Participants with Excluded Subscales due to Missing Data

	Wave 1	Wave 2	Wave 3
# of participants with excluded subscales	27	15	7
# of participants missing the Social-Emotional subscale	10	4	1
# of participants missing the Physical subscale	7	2	2
# of participants missing the Language subscale	12	4	2
# of participants missing the Cognitive subscale	17	9	2
# of participants missing the Literacy subscale	21	13	7
# of participants missing the Mathematics subscale	21	15	5

Data Analysis

Given the clustered nature of the data collected for this evaluation (e.g., children within classrooms, teachers within centers), traditional data analysis methods, which assume data from all participants are independent of each other, could not be used. Rather, analysis techniques that take into account the dependence among observations (e.g., children within a classroom are more similar to each other than are children from different classrooms) were used. In particular, the goal was to accurately estimate the relationships among relevant predictors (e.g., ITERS) and outcomes (e.g., TS Gold cognitive development norms), controlling for other substantive variables (see next section on covariates), so a group of analytic models known as population average methods (PAMS; McNeish et al., 2017) were employed using the statistical software packages SAS 9.4 and SPSS 27. When necessary, adjustments were applied to the analyses to correct issues that can arise from using PAMS with small sample sizes (McNeish & Harring, 2017).

Originally, the goal of the analyses was to model relationships among variables across time. However, several limitations of the data prevented the evaluation team from pursuing these analyses. These limitations include (1) lack of variability in scores for some outcome variables across time, (2) insufficient sample sizes for the model complexity inherent in the longitudinal analyses, (3) issues with missing data across time, and (4) the unmeasured influence of the COVID-19 pandemic that began approximately halfway through the evaluation. Thus, all analyses were conducted separately by wave. Whenever possible, predictor variables and covariates measured at the specific wave under investigation were included in the analyses. When this was not possible, the preceding wave's values were used, and, in instances where the variable was only collected at one wave, the values from that wave were used.

Covariates

In this evaluation, the evaluation team examined whether certain characteristics of early childhood education were related to child outcomes. In these analyses, accounting or controlling for other variables (i.e., covariates) that can affect the outcomes of interest was essential. These variables could be established theoretically or using basic descriptive statistical analyses. Using theoretical considerations and examining correlations within the present sample, the evaluation team identified 11 covariates for the analyses of the child outcomes: parent perceived stress, family-work conflict, child gender, child race, child ethnicity, child disability or special needs status, parent education, number of family moves, number of parental deployments, hours per week in child care, and the occurrence of major life changes in the preceding months. However, subsequent analyses found that family-work conflict, number of parental deployments, and hours per week in child care were missing data for a substantial number of participants. Therefore, to retain as many participants as possible in the analyses, those covariates were removed from final analyses.

Similarly, the evaluation team identified nine covariates for the analyses of the parent outcomes: child disability or special needs status, years attending the child care center, hours per week in child care, respondent gender, dual military status, parental time away from child in previous year,

number of parents who spent 30 days or more away from their child in the last year, the occurrence of major life changes in the preceding months, and the number of family relocations. Subsequent analyses found that hours per week in child care, parental time away from child in previous year, and number of parents who spent 30 days or more away from their child in the last year were missing data for a substantial number of participants and, therefore, were removed from the final analyses.

Results

Demographic, Process, and Descriptive Data

Parent Demographic Data

The majority of respondents to the parent questionnaire were mothers (i.e., 86-88%, depending on wave). In Waves 1 and 2, one grandparent completed the questionnaire. The majority of respondents were married for the first time (i.e., 68-70%) and between 30 and 39 years old at Wave 1 (i.e., 59%). Figures 9, 10, and 11 illustrate the breakdown of respondents' relationship to the child, respondents' relationship status, and respondents' ages.

Figure 9
Respondents' Relationship to Child

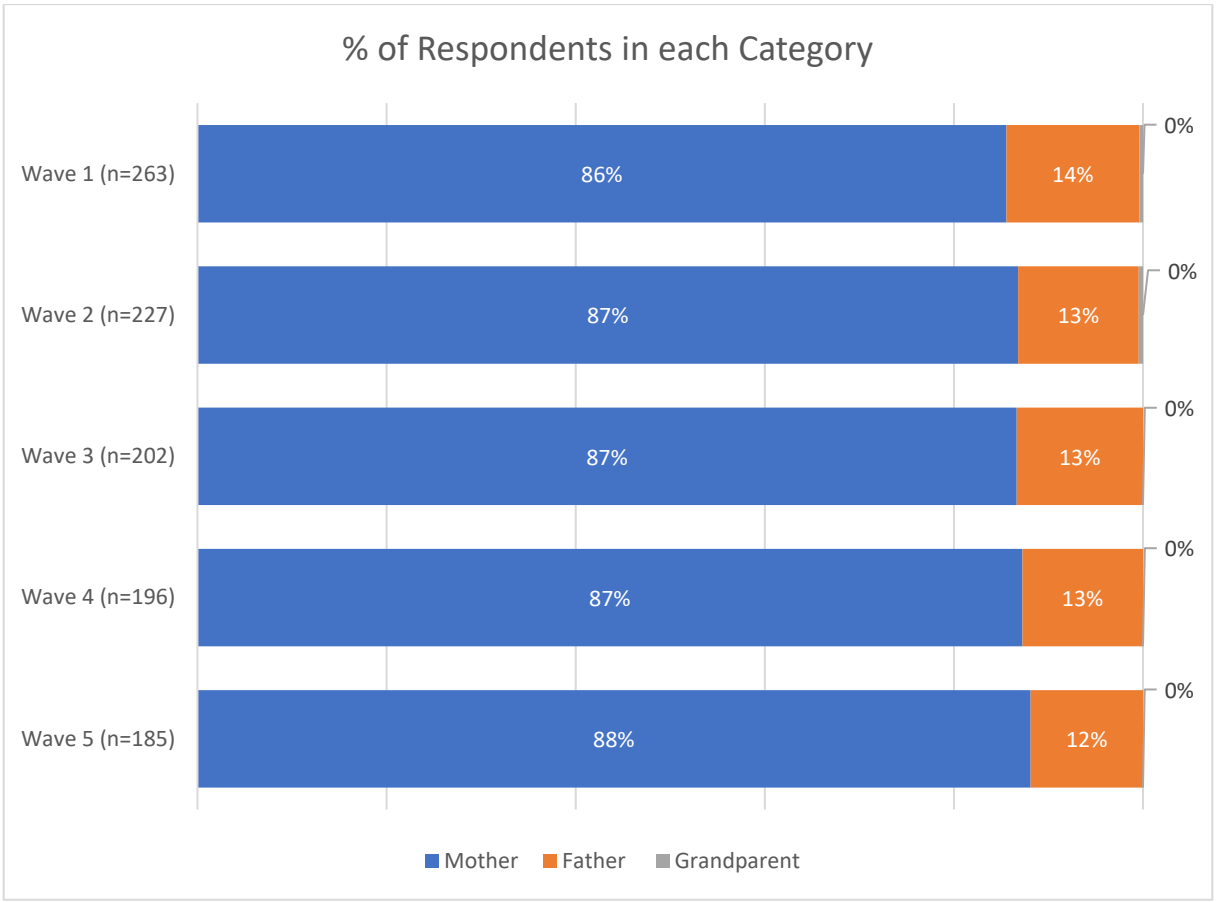


Figure 10
Respondents' Marital Status

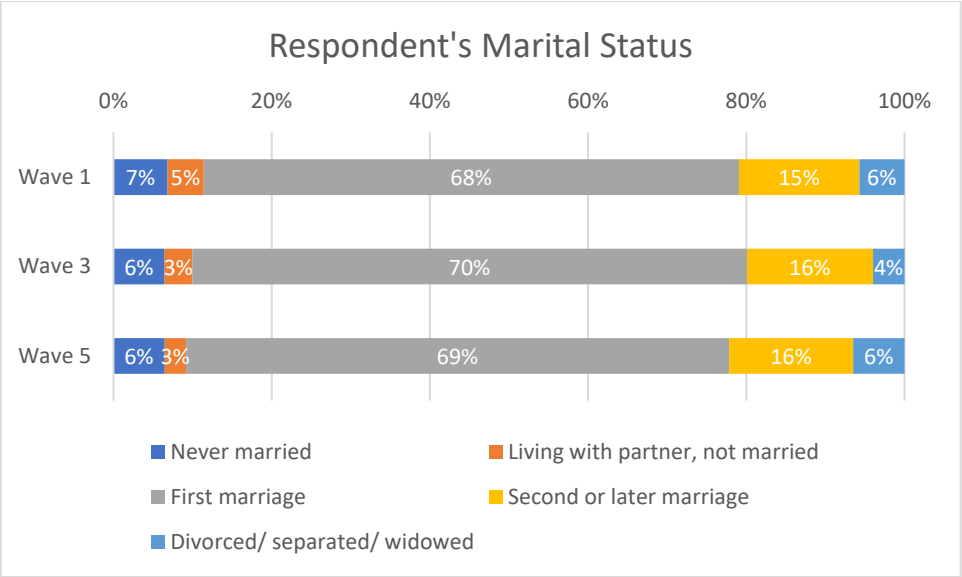
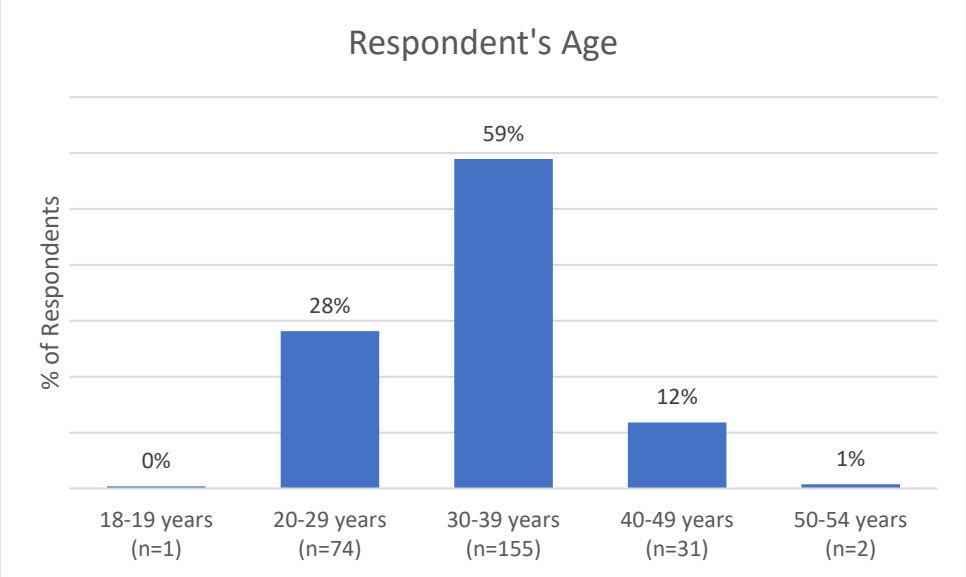


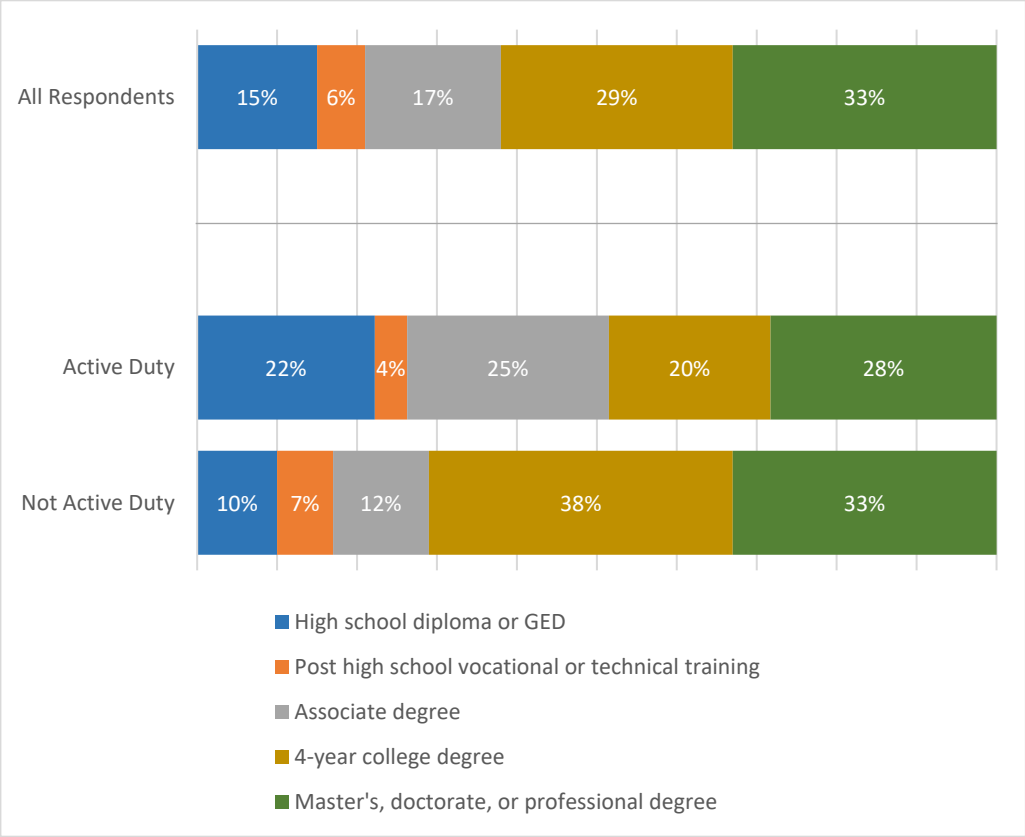
Figure 11
Respondents' Age Range at Wave 1



Overall, respondents were highly educated. Respondents' education levels are shown in Figure 12 for the whole sample and are broken down by whether the respondent was active duty military or not. Twenty-two participants are included in the *All Respondents* bar but are not included in the *Active Duty* or *Not Active Duty* bars because they either did not indicate a military affiliation or their responses were unclear. The *Not Active Duty* category could be any combination of the

following: military spouses, DoD civilian employees, National Guard/Reserve members, and veterans.

Figure 12
Respondents' Highest Level of Education at Wave 1 for the Whole Sample and Separated by Military Affiliation



Respondents could select among six options to describe their military affiliation and their child's other parent's military affiliation. They were asked to select all that apply. Table 18 displays the number and percentage of respondents who selected each option. Because respondents could select more than one option, the total number is more than the number of respondents, and the percent adds up to more than 100%.

Table 18
Military Affiliation of the Respondent and the Child's Other Parent's Military Affiliation

Affiliation - Respondent	Wave 1 # (%)	Wave 3 # (%)	Wave 4 # (%)	Wave 5 # (%)
Respondent				
Civilian - DoD employee	73 (30%)	58 (29%)	61 (31%)	60 (32%)
Civilian - non-DoD employee	48 (20%)	47 (23%)	45 (23%)	48 (26%)
Currently serving - Active Duty	103 (43%)	80 (40%)	72 (37%)	60 (32%)
Currently serving - Reserve or National Guard	6 (2%)	7 (4%)	7 (4%)	5 (3%)
Veteran - Active Duty	25 (10%)	14 (7%)	23 (12%)	19 (10%)
Veteran - Reserve or National Guard	8 (3%)	1 (1%)	3 (2%)	4 (2%)
Child's Other Parent				
Civilian - DoD employee	28 (12%)	21 (10%)	18 (9%)	19 (10%)
Civilian - non-DoD employee	40 (17%)	31 (15%)	30 (15%)	32 (17%)
Currently serving - Active Duty	147 (61%)	113 (56%)	103 (53%)	93 (50%)
Currently serving - Reserve or National Guard	5 (2%)	3 (2%)	2 (1%)	5 (3%)
Veteran - Active Duty	33 (14%)	26 (13%)	31 (16%)	31 (17%)
Veteran - Reserve or National Guard	4 (2%)	5 (3%)	5 (3%)	4 (2%)
Dual Military	59 (24%)	49 (24%)	45 (23%)	35 (19%)

Tables 19, 20, and 21 present the Service(s) with which the respondent and the child's other parent were affiliated, the number of years the respondent and other parent had been in the Service, and the current paygrade for the individuals who are currently serving. Changes across time could be due to changes in status or due to attrition.

Table 19
The Service Affiliation of Currently Serving Parents

Service (current Service member)	Wave 1 # (%)	Wave 3 # (%)	Wave 5 # (%)
Respondent			
Army	43 (41%)	33 (40%)	25 (40%)
Marine Corps	14 (13%)	14 (17%)	8 (13%)
Navy	49 (46%)	35 (43%)	30 (48%)
Air Force	0 (0%)	0 (0%)	0 (0%)
Child's Other Parent			
Army	52 (34%)	33 (29%)	28 (29%)
Marine Corps	37 (25%)	32 (28%)	26 (27%)
Navy	54 (36%)	42 (37%)	38 (39%)
Air Force	8 (5%)	6 (5%)	5 (5%)

Table 20
The Average and Range of Years that Currently Serving Parents Have Served in the Military

Years in the Military (current Service member)	Wave 1	Wave 3	Wave 5
Respondent			
Average	10 years	11 years	12 years
Range	1 - 24 years	2 - 25 years	3 – 25 years
Child's Other Parent			
Average	11 years	12 years	12 years
Range	1 - 28 years	2 – 26 years	2 – 22 years

Table 21
The Paygrade of Currently Serving Parents

Paygrade (current Service member)	Wave 1 # (%)	Wave 3 # (%)	Wave 5 # (%)
Respondent			
E1-E2	0 (0%)	0 (0%)	0 (%)
E3-E4	16 (15%)	8 (9%)	5 (8%)
E5-E6	43 (39%)	33 (38%)	24 (38%)
E7-E9	17 (16%)	16 (19%)	12 (19%)
W1-W2	0 (0%)	0 (0%)	0 (0%)
W3-W5	2 (2%)	2 (2%)	2 (3%)
O1-O2	1 (1%)	1 (1%)	0 (0%)
O3-O4	23 (21%)	20 (23%)	15 (23%)
O5-O6	7 (6%)	6 (7%)	6 (9%)
Child's Other Parent			
E1-E2	1 (1%)	0 (0%)	0 (0%)
E3-E4	9 (6%)	6 (5%)	6 (6%)
E5-E6	59 (39%)	45 (39%)	42 (43%)
E7-E9	38 (25%)	28 (24%)	23 (24%)
W1-W2	1 (1%)	1 (1%)	1 (1%)
W3-W5	3 (2%)	2 (2%)	2 (2%)
O1-O2	3 (2%)	3 (3%)	1 (1%)
O3-O4	31 (21%)	25 (22%)	16 (17%)
O5-O6	6 (4%)	6 (5%)	6 (6%)

For parents who are veterans, Tables 22, 23, and 24 show the Service with which the respondent and the child's other parent were affiliated, the number of years the respondent and the child's other parent served, and the paygrade at separation from military service.

Table 22
The Service Affiliation of Veteran Parents as Reported at Wave 1

Service (veterans)	Respondent # (%)	Child's Other Parent # (%)
Army	15 (43%)	14 (50%)
Marine Corps	10 (29%)	4 (14%)
Navy	7 (20%)	7 (25%)
Air Force	3 (9%)	3 (11%)

Table 23
The Average and Range of Years that Veteran Parents Served in the Military as Reported at Wave 1

Years in the Military (veterans)	Respondent	Child's Other Parent
Average	8 years	9 years
Range	1 – 23 years	2 - 23 years

Table 24
The Paygrade of Veteran Parents as Reported at Wave 1

Paygrade Last Day of Service (veterans)	Respondent # (%)	Child's Other Parent # (%)
E1-E2	0 (0%)	1 (3%)
E3-E4	7 (20%)	11 (30%)
E5-E6	20 (57%)	16 (43%)
E7-E9	4 (11%)	5 (14%)
W1-W2	0 (0%)	0 (0%)
W3-W5	0 (0%)	1 (3%)
O1-O2	0 (0%)	1 (3%)
O3-O4	3 (9%)	2 (5%)
O5-O6	1 (3%)	0 (0%)

If the respondents indicated that they or their child's other parent was a Service member (i.e., currently or a veteran), they were asked how many deployments the Service member parent(s) had during the child's life. The number of deployments reported by the respondent, for him or herself and for the child's other parent, ranged from 0 to 5 deployments. At Wave 1, the majority of the children (i.e., 56%) had not experienced a parental deployment. At Wave 5, 50% of the children had not experienced a parental deployment. The change from Wave 1 to Wave 5 could be due to attrition or due to deployments that occurred during the evaluation.

The number of deployments that occur during a child's life depends, in part, on the child's age. For families in which the respondent was a Service member, Figure 13 shows the number of respondent deployments the family experienced during the child's life, separated by the child's age range. For families in which the child's other parent was the Service member, Figure 14 shows the number of the child's other parent's deployments that the family experienced during the child's life, separated by the child's age range.

Figure 13
The Number of Respondent Deployments During the Child's Life – by Child Age Group

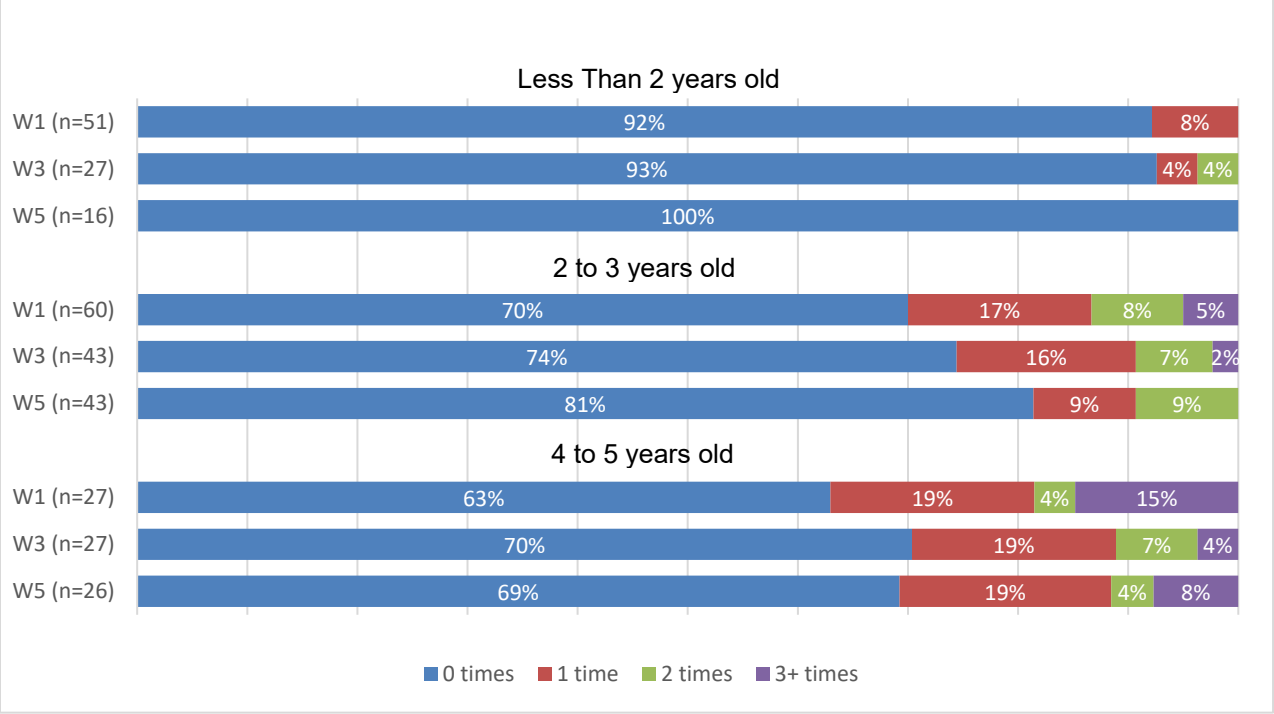


Figure 14
The Number of Deployments for the Child's Other Parent During the Child's Life – by Child Age Group

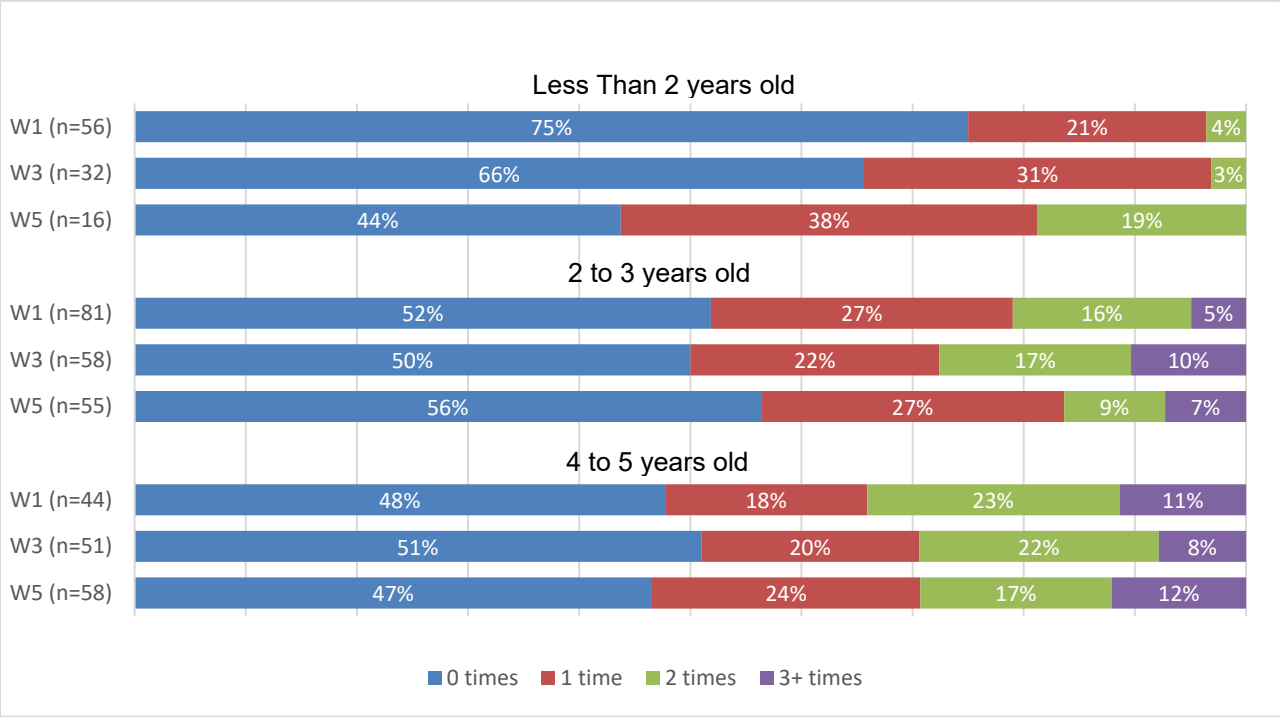


Figure 15 shows the total number of parental (i.e., the respondent and the child's other parent) deployments during the child's life. As shown in this figure, most children (n=123) did not experience any parental deployments. Ninety-six children experienced at least one parental deployment, and two children experienced seven parental deployments.

Figure 15
The Total Number of Parental Deployments During the Child's Life – by Child Age Group

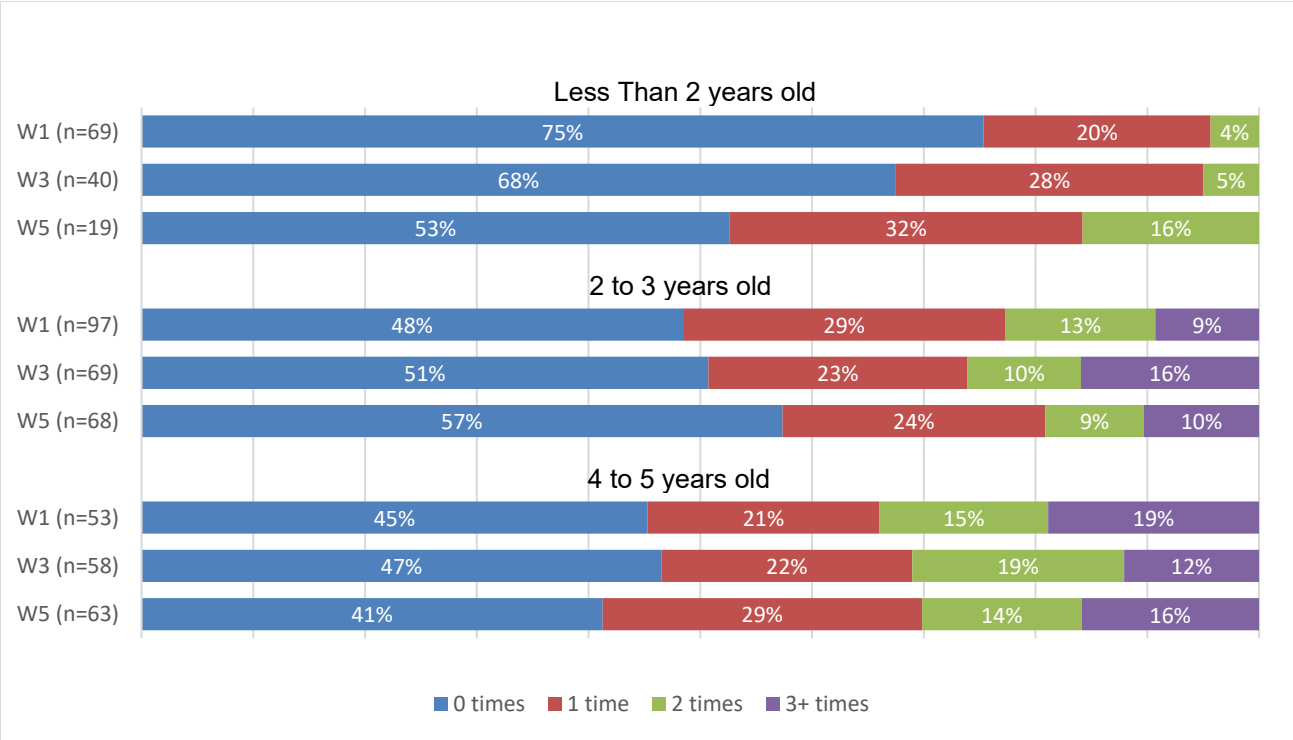


Table 25 illustrates the amount of time the respondent and the child's other parent were away from the child in the last year. This was only asked for parents who were currently serving and could include deployments, trainings, or temporary duty (TDY).

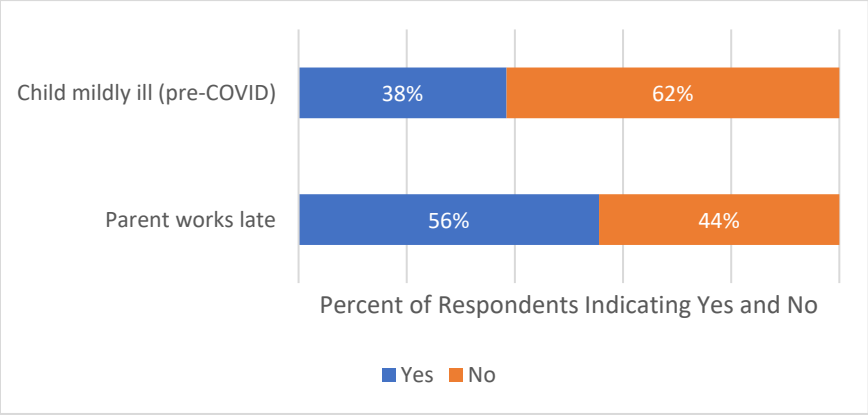
Table 25
The Amount of Time Parents Were Away from Their Children in the Last Year

Time Away from Child in the Last Year Due to Deployment, Training, or TDY (currently serving)	Wave 1 # (%)	Wave 3 # (%)	Wave 5 # (%)
Respondent			
0 days	35 (32%)	27 (31%)	29 (46%)
1-29 days	32 (29%)	33 (38%)	21 (33%)
30 days-6 months	35 (32%)	20 (23%)	13 (21%)
7-12 months	7 (6%)	6 (7%)	0 (0%)
Child's Other Parent			
0 days	13 (8%)	16 (14%)	12 (12%)
1-29 days	46 (30%)	33 (29%)	26 (27%)
30 days-6 months	63 (41%)	41 (36%)	41 (42%)
7-12 months	32 (21%)	25 (22%)	19 (19%)

Child Care Availability

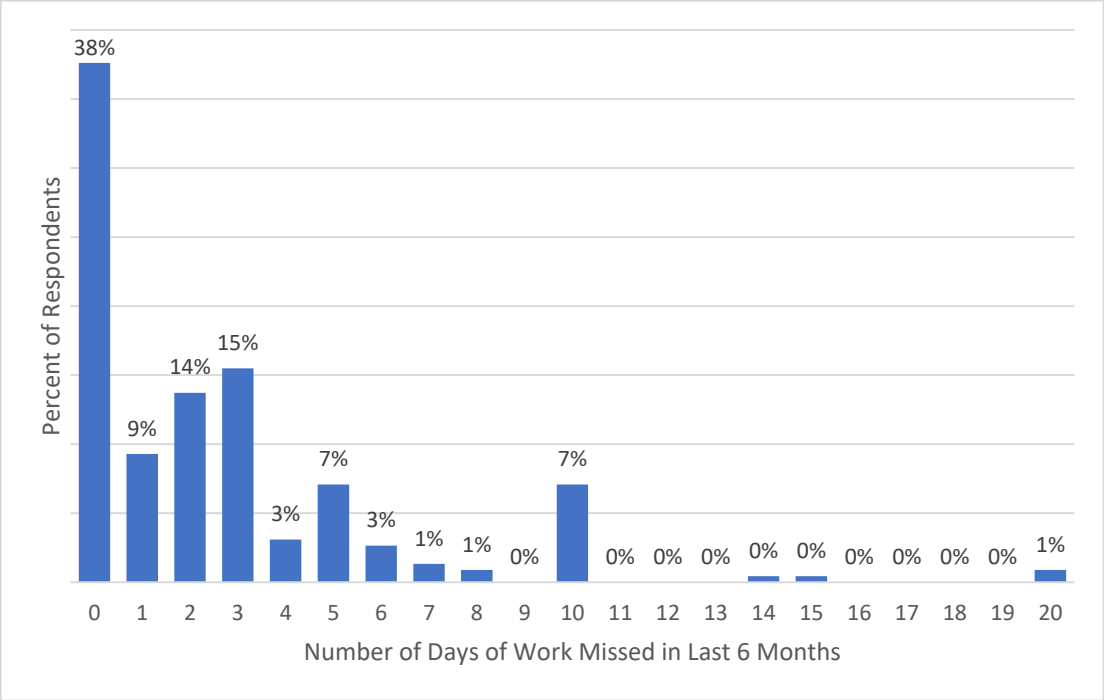
Parents reported on several items that were specifically related to child care. For Wave 1, these questions were asked before the COVID-19 pandemic began. Figure 16 shows the responses to questions about whether child care is available when their child is mildly ill and when the respondent works late. Although a slight majority of parents indicated that child care is available when they work late, the majority indicated that child care is not available when their child is mildly ill.

Figure 16
Child Care Availability when Child is Mildly Ill or When Respondent Works Late (Wave 1)



Responding parents missed an average of 2.63 days of work in the previous 6 months due to child care arrangements. The number of missed days ranged from 0 to 20 days. Although the largest number of people missed between 0 and 3 days of work due to child care arrangements, a non-trivial number (i.e., n=20; 9%) missed 2 to 4 weeks of work during the previous 6 months. Figure 17 displays the number of missed days of work.

Figure 17
The Percent of Respondents Who Missed Work in the Last 6 Months Due to Child Care Arrangements (n=226) (Wave 1)



Child Demographic Data

Children in participating families were 0 to 5 years old. The majority were 2 to 3 years old (see Figure 18). Slightly more than half were male (see Figure 19). Children were White (n=195), Black or African American (n=74), Asian American (n=23), American Indian or Alaska Native (n=6), and Native Hawaiian or other Pacific Islander (n=4). The number of children in each race category adds up to more children than there were in participating families. This is because, for 40 children, respondents endorsed more than one category. In Figure 20, race categories are mutually exclusive, and all children whose parents selected more than one race category are represented in the multiracial category. As shown in Figure 21, 23% of children are Spanish, Hispanic, or Latino.

Figure 18
Children's Age

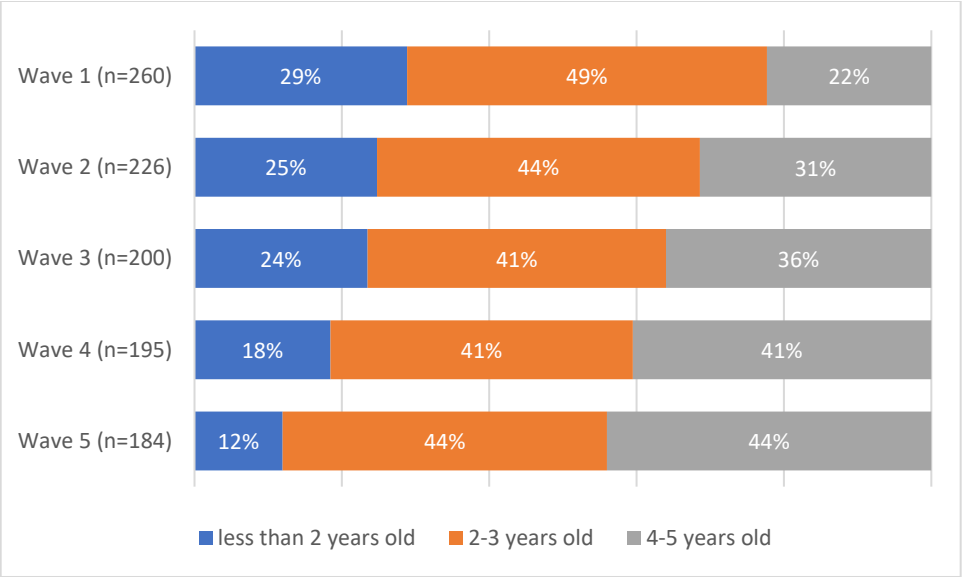


Figure 19
Children's Gender (Wave 1)

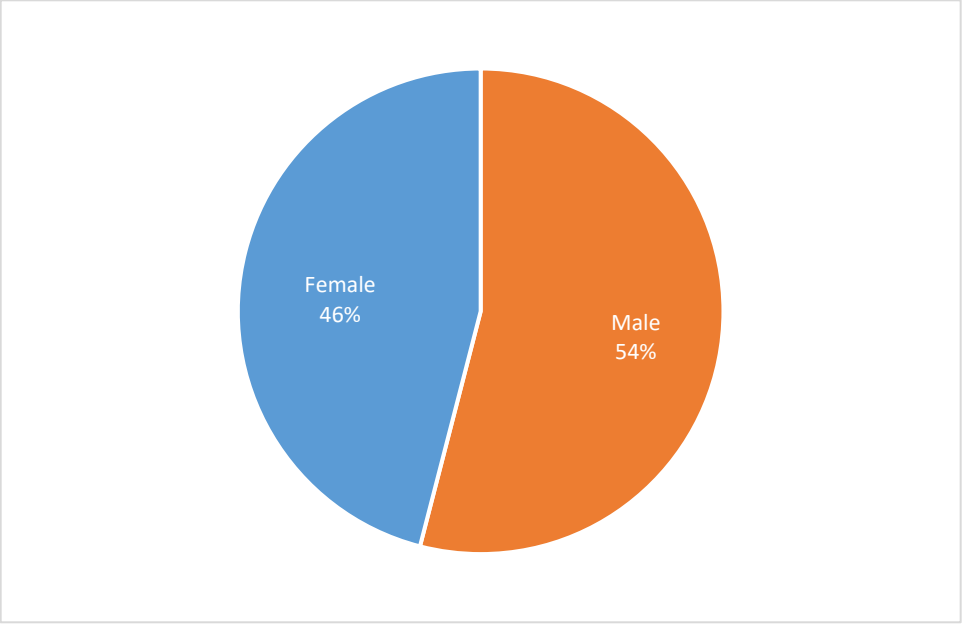


Figure 20
Children's Race (Wave 1)

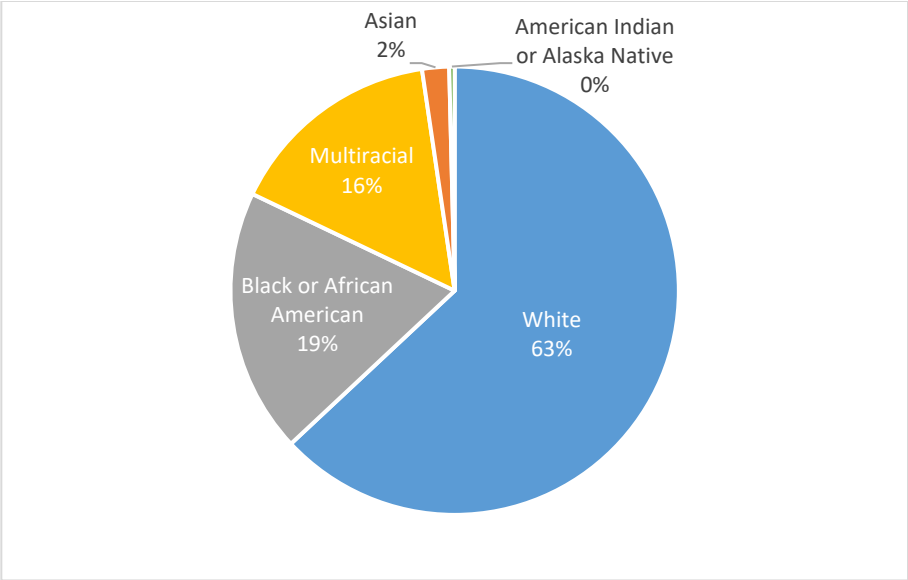
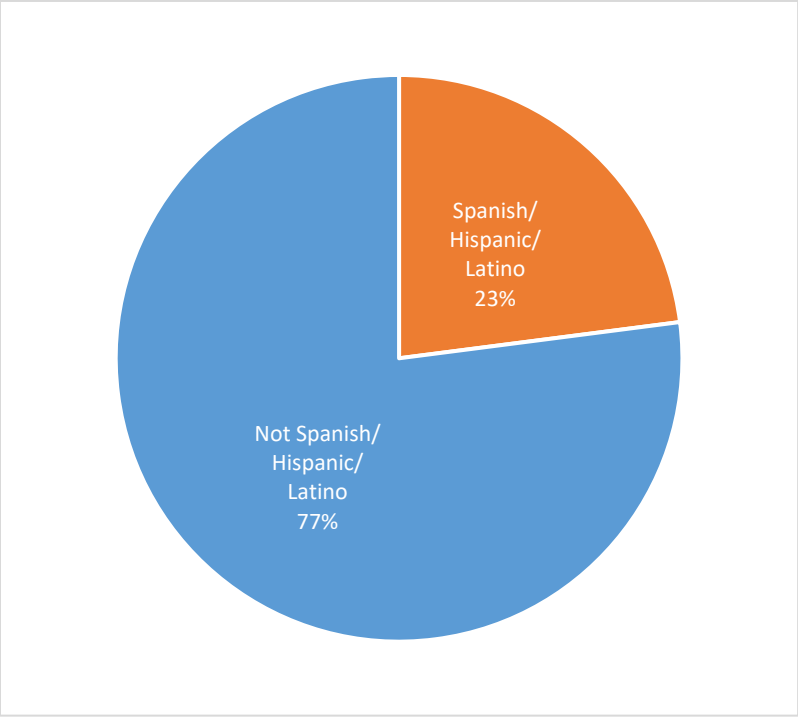
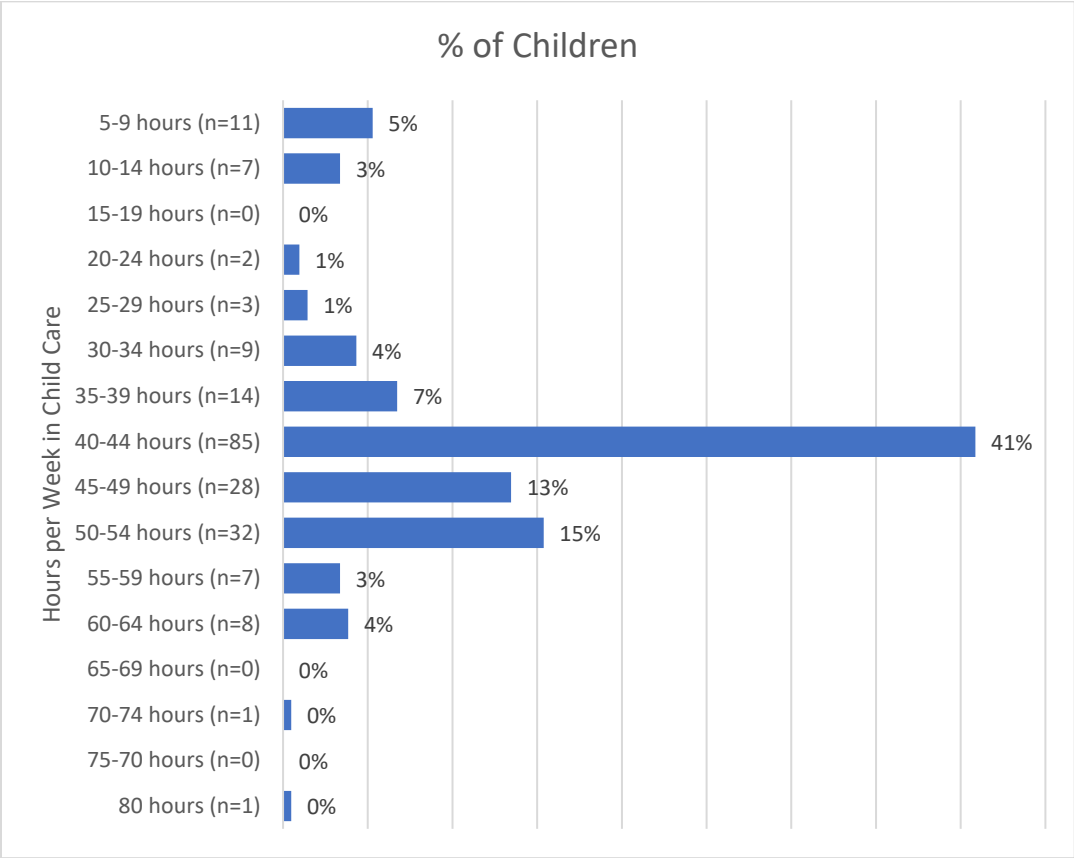


Figure 21
Children's Ethnicity (Wave 1)



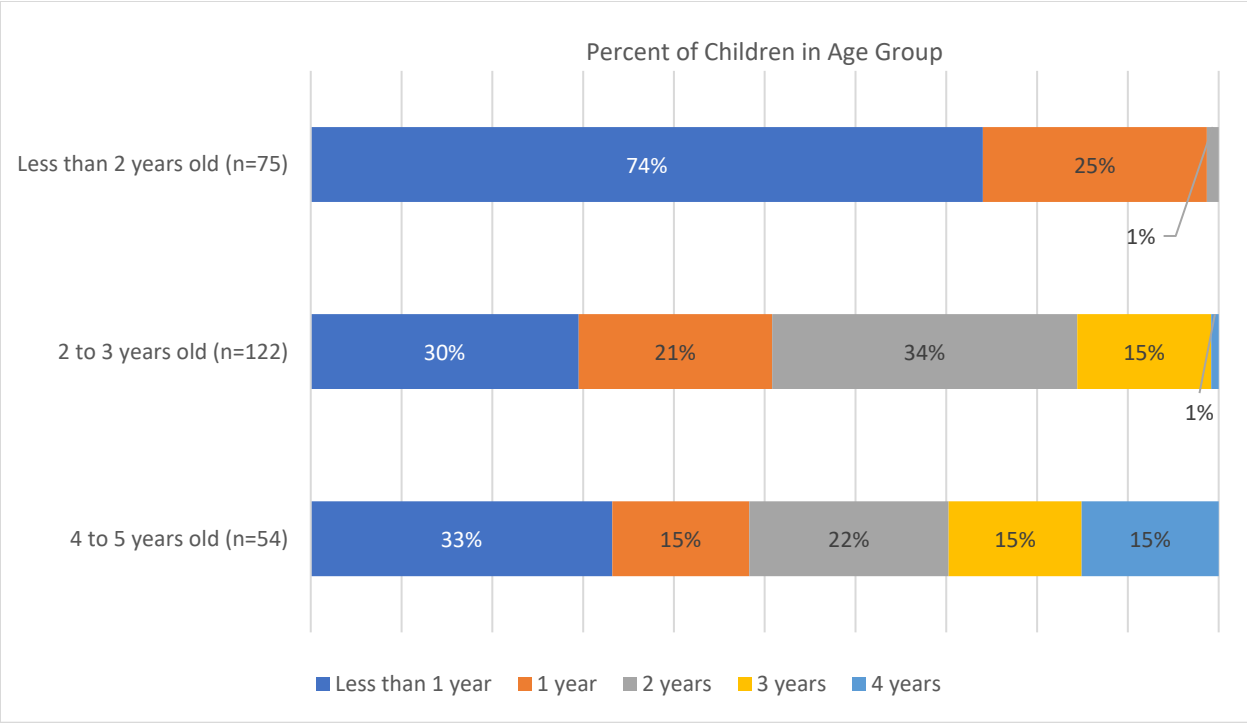
Children spend an average of 40.16 hours per week in child care, but the range was from 5 to 80 hours. Figure 22 shows the number of hours per week children spend in child care in 5-hour increments.

Figure 22
The Number of Hours Children Spend in Child Care per Week (Wave 1)



The length of time that children have been at their current child care center is represented in Figure 23. Because the length of time in the center is, in part, dependent on the child's age, in the figure, the length of time at the center is split by child age range.

Figure 23
Length of Time Children Have Attended Their Current Child Care Center (Wave 1)



Center Type Differences in Demographic Variables

A few differences emerged between the CDC participants and the civilian center participants in relation to the demographic and military-related variables (see Table 26). Parent respondents in the civilian centers were more likely to be fathers, DoD civilian employees, or to have a paygrade of E7-E9 or W3-W5 (the child's other parent was more likely to have a paygrade of O3-O4); were less likely to be dual military or active duty; and, on average, were older, had spent more time in the military, and had spent more time away from their child due to deployments, TDY, or training. Differences also emerged for the number of deployments the child experienced, but these differences varied by child age and wave of data collection.

Table 26

Differences Between Participants in CDCs and Civilian Centers on Parent-Reported Demographic Variables

	CDC	Civilian Centers	No Significant Difference
Respondent Relationship to Child	W3: ↓ % of father respondents	W3: ↑ % of father respondents	W1, W2, W4, W5
Parent Age	W1, W3: Younger	W1, W3: Older	W5
Parent Relationship Status			W1, W3, W5
Dual Military	W1 & W4: ↑ % dual military	W1 & W4: ↓ % dual military	W3, W5
Civilian DoD Employee (respondent)	W1, W3, W4, W5: ↓ % DoD civilian employee	W1, W3, W4, W5: ↑ % DoD civilian employee	
Active Duty (respondent)	W1 & W4: ↑ % active duty	W1 & W4: ↓ % active duty	W3, W5
National Guard/Reserves (respondent)			W1, W3, W4, W5
Active Duty Veteran (respondent)			W1, W3, W4, W5
National Guard/Reserves Veteran (respondent)			W1, W3, W4, W5
Civilian DoD Employee (other parent)			W1, W3, W4, W5
Active Duty (other parent)			W1, W3, W4, W5
National Guard/Reserves (other parent)			W1, W3, W4, W5
Active Duty Veteran (other parent)			W1, W3, W4, W5
National Guard/Reserves Veteran (other parent)			W1, W3, W4, W5
Child Care Available when Child Mildly Ill			W1, W3, W4, W5
Child Care Available when Working Late			W1, W3, W4, W5
Education (respondent)			W1, W3, W5
Paygrade (respondent)	W1, W3, W5: ↓ % of E7-E9 W1 & W5: ↓ % of W3-W5	W1, W3, W5: ↑ % of E7-E9 W1 & W5: ↑ % of W3-W5	
Paygrade (other parent)	W1: ↓ % of O3-O4	W1: ↑ % of O3-O4	W3, W5
Child Age Category			W1-W5
Years at Current Center			W1, W3, W5
Time in the Military (respondent)	W1, W3, W5: Less time	W1, W3, W5: More time	
Time in the Military (other parent)			W1, W3, W5
# of Deployments (respondent)	W3 (0-1 y/o): ↓ deployments	W3 (0-1 y/o): ↑ deployments	W1, W3 2-5 y/o, W5
# of Deployments (other parent)			W1, W3. W5
# of Deployments (total family)			W1, W3. W5
Time Away (respondent)	W3 less time away	W3 more time away	W1, W5
Time Away (other parent)			W1, W3, W5
Hours in Child Care			W1, W3, W5

Note. Significant differences listed are at the $p < .05$ level. Only # of Deployments was split by child age group for this analysis. y/o = years old. W1, W2, W3, W4, W5 indicates, respectively, Wave 1, Wave 2, Wave 3, Wave 4, and Wave 5.

Leadership Support Descriptive Data

Directors

CDC directors reported on their perceptions of leadership support. There was little variability in director responses. All received responses were either “Agree (4)” or “Strongly Agree (5).” Table 27 shows the minimum and maximum responses and the average and standard deviation across all respondents. Because there was very little variability in the results, these variables were not able to be used in subsequent outcome analyses.

Table 27
Director Report Descriptive Statistics for Leadership Support Questions

	Minimum	Maximum	Average	Standard Deviation
Command at this installation is supportive of the CDC.	4	5	4.71	0.49
Command at this installation is supportive of my role as director.	4	5	4.71	0.49

Direct-Care Staff

Direct-care staff also reported on their perception of leadership support. Responses spanned the entire range of response options (i.e., 1 = *Strongly Disagree*, 2 = *Disagree*, 3 = *Neither Agree Nor Disagree*, 4 = *Agree*, 5 = *Strongly Agree*). Table 28 shows the minimum and maximum responses and the average and standard deviation across all respondents.

Table 28
Direct-Care Staff Descriptive Statistics for Leadership Support Questions

	Minimum	Maximum	Average	Standard Deviation
The director is effective in addressing the needs of the direct-care staff. (n=118)	1	5	3.66	1.29
The director is effective in addressing the needs of enrolled children and families. (n=117)	1	5	4.01	1.08
Command at this installation is supportive of the CDC. (n=114)	1	5	3.76	1.18

Item-level responses for direct-care staff report of leadership support are shown in Table 29.

Table 29
Direct-Care Staff Item-level Frequencies

	Strongly Disagree n (%)	Disagree n (%)	Neither Agree nor Disagree n (%)	Agree n (%)	Strongly Agree n (%)
The director is effective in addressing the needs of the direct-care staff. (n=118)	12 (10%)	12 (10%)	16 (14%)	42 (36%)	36 (31%)
The director is effective in addressing the needs of enrolled children and families. (n=117)	7 (6%)	3 (3%)	16 (14%)	47 (40%)	44 (38%)
Command at this installation is supportive of the CDC. (n=114)	9 (8%)	5 (4%)	26 (23%)	38 (33%)	36 (32%)

Note: Due to rounding, totals may not equal 100%.

As can be seen in the item-level analyses, a lot of variability in the direct-care staff responses for the leadership support questions exists. There was some clustering of responses by center. Tables 30, 31, and 32 display center-level analyses for each question.

Table 30
Center Level Analyses: The Director is Effective in Addressing the Needs of the Direct-Care Staff

	# of Centers (out of 12)
Greater than 80% of responding direct-care staff at the center responded Agree or Strongly Agree	5
At least 25% of responding direct-care staff at the center responded Disagree or Strongly Disagree	5

Table 31

Center Level Analyses: The Director is Effective in Addressing the Needs of Enrolled Children and Families

	# of Centers (out of 12)
Greater than 80% of responding direct-care staff at the center responded Agree or Strongly Agree	6
At least 25% of responding direct-care staff at the center responded Disagree or Strongly Disagree	1

Table 32

Center Level Analyses: Command at the Installation is Supportive of the CDC

	# of Centers (out of 12)
Greater than 80% of responding direct-care staff at the center responded Agree or Strongly Agree	3
Greater than 25% of responding direct-care staff at the center responded Disagree or Strongly Disagree	1

Classroom Characteristics

Direct-Care Staff Education

In 2020, the NAEYC updated their standards. Included in this update is a description of the education expectations for assistant teachers/teacher aides and for teachers (NAEYC, 2019). Current expectations are that assistant teachers/teacher aides have a minimum of a Child Development Associate (CDA) certificate or 12 college credits in early childhood education, child development, elementary education, and/or early childhood special education. Current expectations for teachers can be achieved in one of three ways: (1) a minimum of an associate's or bachelor's degree in early childhood education, child development, elementary education, or early childhood special education; (2) a degree in a different major with at least 36 college credits in early childhood education, child development, elementary education, and/or early childhood special education; or (3) a state public school teacher certification for age birth to 8 years.

Because measures for this evaluation were developed before the new standards were released, the measures employed in this evaluation do not align with current NAEYC teacher education standards. Therefore, when coding the education level of the direct-care staff, the categories were created to best match current NAEYC standards, taking into account the manner in which the questions were asked.

The CEQ requested education and experience information be provided for all direct-care staff in the designated classroom. It provided space for responses for four direct-care staff members. However, responses did not appear to be provided for all direct-care staff members in each classroom. When comparing the number of direct-care staff who were in the room during the observation with the number of direct-care staff for whom responses were provided on the CEQ, the numbers did not always match. For 21 classrooms, the number of direct-care staff with responses on the CEQ matched the number of direct-care staff who were present at the observation. For 51 classrooms, the number of direct-care staff with responses on the CEQ exceeded the number of direct-care staff who were present at the observation. This seems reasonable as not all direct-care staff would necessarily have been present during the observation. However, for 50 classrooms, the number of direct-care staff present during the observation exceeded the number of direct-care staff with responses on the CEQ. This suggests that complete information was not received on the CEQ. Had the COVID 19 pandemic not occurred and the evaluation team had been able to collect these data at Waves 3 and 5, the team would have taken steps to obtain more complete information.

Table 33 shows the direct-care staff education categorization and the number of CDC classrooms that fall into each category, based on the information that was provided. The number of classrooms that did not report a direct-care staff member with a bachelor's degree (n=101) far exceeds those that did report a direct-care staff member with a bachelor's degree (n=21). The small number of classrooms at the higher education levels will negatively impact subsequent analyses that examine the influence of direct-care staff education on classroom quality and child outcomes.

Table 33

The Number of Classrooms in Each Education Category

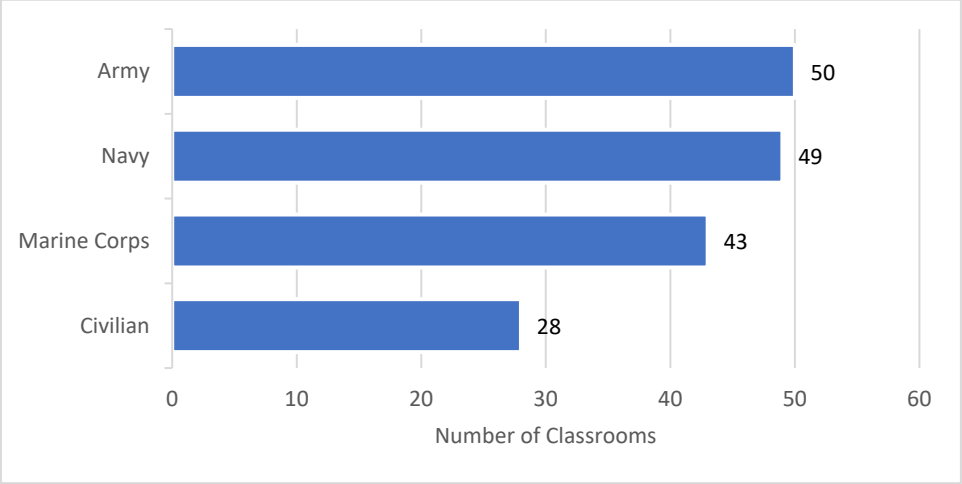
Number of Classrooms	Direct-Care Staff Education Category
1	At least one individual with a bachelor's degree; all others have a certificate/associate's degree
5	At least one individual with a bachelor's degree; some others have a certificate/associate's degree
15	At least one individual with a bachelor's degree; no others have a certificate/associate's degree
22	No individual with a bachelor's degree; all others have a certificate/associate's degree
37	No individual with a bachelor's degree; some others have a certificate/associate's degree
42	No individual with a bachelor's degree; no others have a certificate/associate's degree

Classroom Quality

Between August 2019 and March 2020, 170 classroom observations were conducted by independent observers. Figure 24 shows the number of observations conducted for each Service and the civilian centers.

Figure 24

Number of Classrooms Observed

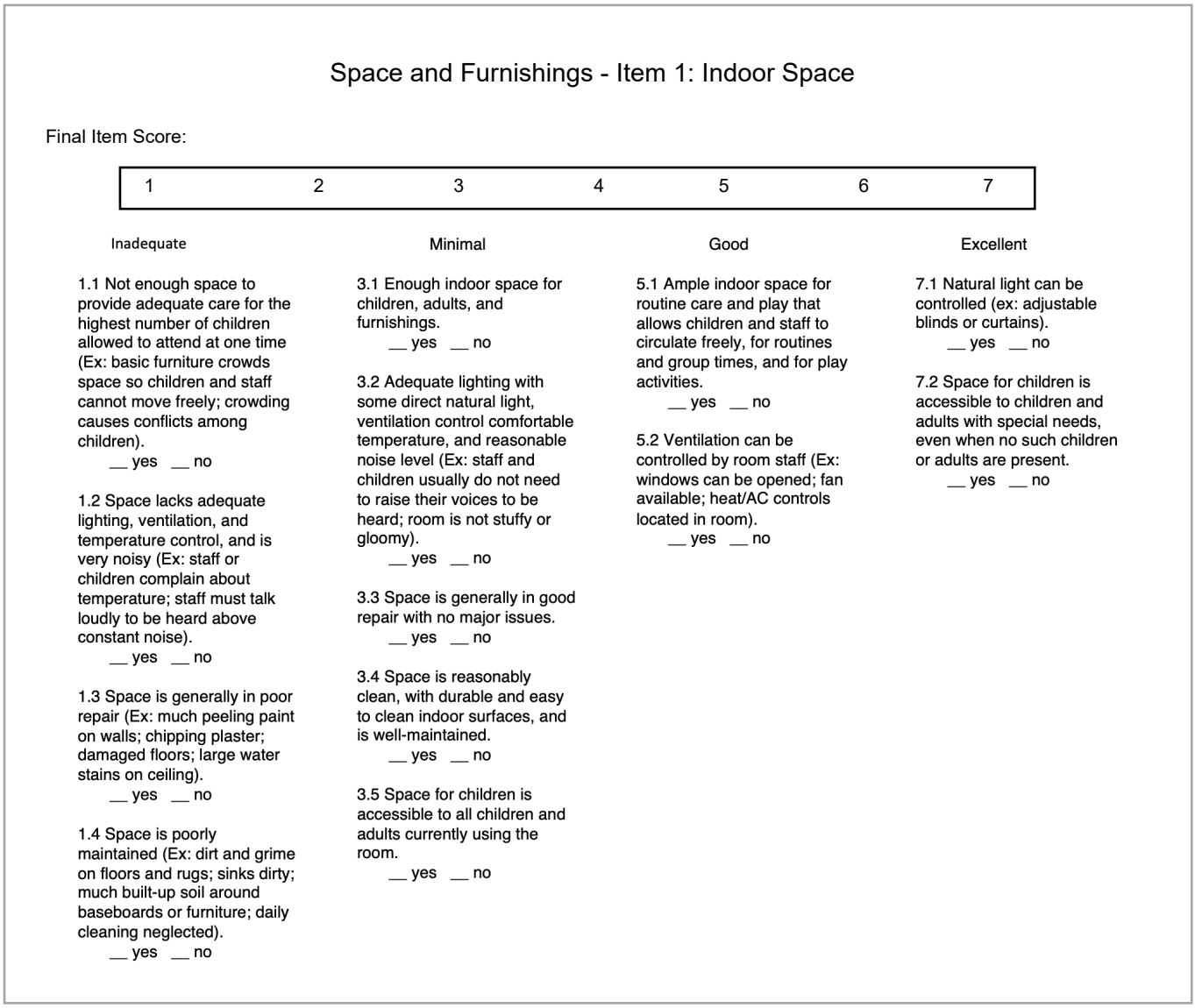


The ITERS and the ECERS are widely used assessments that measure the quality of child care center classrooms. Items fall into six subscales: Space and Furnishing, Personal Care Routines, Language and Books/Literacy, Activities/Learning Activities, Interaction, and Program Structure. Each item is comprised of indicators - specific characteristics of the classroom or interactions that receive either a *yes* or a *no* score. These indicators fall into four quality categories: Inadequate (1), Minimal (3), Good (5), and Excellent (7). The indicator scores are used to create the item score: Less than Minimal (1-2), Minimal (3-4), Good (5-6), and Excellent (7).

In order to score an Excellent, the classroom must have a satisfactory score on every indicator, at every level, that comprises that item (i.e., a *no* on each of the inadequate-level items and a *yes* on each of the minimal-, good- and excellent-level items). In order to score Less Than Minimal, only one poor score (i.e., a *yes* on the inadequate-level items or a *no* on the minimal-level items) on an inadequate- or minimal-level indicator is needed. The item scores are averaged to create the subscale score. See an example item in Figure 25.

Figure 25

Example Item and Associated Indicators from the ITERS



Space and Furnishings - Item 1: Indoor Space

Final Item Score:

1	2	3	4	5	6	7

Inadequate

1.1 Not enough space to provide adequate care for the highest number of children allowed to attend at one time (Ex: basic furniture crowds space so children and staff cannot move freely; crowding causes conflicts among children).
__ yes __ no

1.2 Space lacks adequate lighting, ventilation, and temperature control, and is very noisy (Ex: staff or children complain about temperature; staff must talk loudly to be heard above constant noise).
__ yes __ no

1.3 Space is generally in poor repair (Ex: much peeling paint on walls; chipping plaster; damaged floors; large water stains on ceiling).
__ yes __ no

1.4 Space is poorly maintained (Ex: dirt and grime on floors and rugs; sinks dirty; much built-up soil around baseboards or furniture; daily cleaning neglected).
__ yes __ no

Minimal

3.1 Enough indoor space for children, adults, and furnishings.
__ yes __ no

3.2 Adequate lighting with some direct natural light, ventilation control comfortable temperature, and reasonable noise level (Ex: staff and children usually do not need to raise their voices to be heard; room is not stuffy or gloomy).
__ yes __ no

3.3 Space is generally in good repair with no major issues.
__ yes __ no

3.4 Space is reasonably clean, with durable and easy to clean indoor surfaces, and is well-maintained.
__ yes __ no

3.5 Space for children is accessible to all children and adults currently using the room.
__ yes __ no

Good

5.1 Ample indoor space for routine care and play that allows children and staff to circulate freely, for routines and group times, and for play activities.
__ yes __ no

5.2 Ventilation can be controlled by room staff (Ex: windows can be opened; fan available; heat/AC controls located in room).
__ yes __ no

Excellent

7.1 Natural light can be controlled (ex: adjustable blinds or curtains).
__ yes __ no

7.2 Space for children is accessible to children and adults with special needs, even when no such children or adults are present.
__ yes __ no

The assessment is designed as a time sample, which means that assessments only occur for a specific amount of time. Each classroom is measured for 3 consecutive hours in the morning of 1 day. As such, anomalies may occur where the typical classroom environment may not be represented. Therefore, one should use caution when interpreting instances in which only a few classrooms score especially high or low on a particular item.

In the following two tables, the item scores are used to describe the quality of the CDC classrooms. Table 34 shows the percentage of observed CDC classrooms at each score level (i.e., Excellent, Good, Minimal, Less Than Minimal) on each of the ITERS items. Table 35 illustrates the same information but for the ECERS. Due to rounding, percentages do not always

add up to 100%. In Table 34, there are several items for which at least 25% of the observed CDC classrooms scored an Excellent on ITERS items:

- Indoor space
- Furnishings for care, play, and learning
- Room arrangement
- Safety practices
- Talking with children
- Responding to children's communication
- Encouraging children to communicate
- Gross motor
- Supervision of gross motor play
- Supervision of play and learning (non-gross motor)
- Staff-child interaction
- Providing physical warmth/touch
- Schedule and transitions

Similarly, in Table 35, there are several items for which at least 25% of the observed CDC classrooms scored an Excellent on ECERS items:

- Indoor space
- Furnishings for care, play, and learning
- Room arrangement for play and learning
- Space for privacy
- Space for gross motor play
- Gross motor equipment
- Safety practices
- Individualized teaching and learning
- Staff-child interaction
- Discipline
- Transitions

Table 34

Percentage of Observed CDC Classrooms at Each Score Level on Each ITERS Item (n=98 classrooms)

	Excellent (7)	Good (5-6)	Minimal (3-4)	Less Than Minimal (1-2)	NA
Space and Furnishing					
Indoor space	39%	26%	22%	13%	
Furnishings for care, play, and learning	41%	21%	28%	10%	
Room arrangement	28%	46%	11%	15%	
Display for children	3%	26%	32%	40%	
Personal Care Routines					
Meals/snacks	1%	51%	29%	20%	
Diapering/toileting	10%	14%	71%	4%	
Health practices	5%	4%	16%	74%	
Safety practices	48%	20%	30%	2%	
Language and Books					
Talking with children	51%	20%	19%	9%	
Encouraging vocabulary development	4%	10%	81%	5%	
Responding to children's communication	29%	27%	33%	12%	
Encouraging children to communicate	34%	30%	35%	2%	
Staff use of books with children	13%	35%	30%	22%	
Encouraging children's use of books	3%	33%	38%	27%	
Activities					
Fine motor	12%	30%	37%	21%	
Art	0%	20%	11%	31%	38%
Music and movement	4%	23%	41%	32%	
Blocks	3%	15%	31%	51%	
Dramatic play	1%	28%	18%	53%	
Nature/science	3%	6%	42%	49%	
Math/number	2%	4%	37%	57%	
Appropriate use of technology	0%	0%	3%	3%	94%
Promoting acceptance and diversity	5%	16%	55%	24%	
Gross motor	36%	31%	15%	18%	
Interaction					
Supervision of gross motor play	34%	36%	27%	4%	
Supervision of play and learning (non-gross motor)	35%	21%	34%	10%	
Peer interaction	9%	46%	39%	6%	
Staff-child interaction	48%	15%	27%	10%	
Providing physical warmth/touch	49%	15%	29%	7%	
Guiding children's behavior	10%	31%	42%	17%	
Program Structure					
Schedule and transitions	27%	16%	23%	23%	10%
Free play	14%	52%	28%	6%	
Group play activities	23%	34%	11%	11%	21%

Table 35

Percentage of Observed CDC Classrooms at Each Score Level on Each ECERS Item (n=44 classrooms)

	Excellent (7)	Good (5-6)	Minimal (3-4)	Less Than Minimal (1-2)	NA
Space and Furnishing					
Indoor space	52%	30%	7%	11%	
Furnishings for care, play, and learning	34%	9%	52%	5%	
Room arrangement for play and learning	43%	23%	34%	0%	
Space for privacy	32%	34%	32%	2%	
Child-related display	2%	9%	39%	50%	
Space for gross motor play	52%	16%	27%	5%	
Gross motor equipment	27%	34%	20%	18%	
Personal Care Routines					
Meals/snacks	16%	32%	45%	7%	
Toileting/diapering	20%	11%	61%	7%	
Health practices	5%	18%	45%	32%	
Safety practices	57%	25%	16%	2%	
Language and Literacy					
Helping children expand vocabulary	9%	14%	58%	19%	
Encouraging children to use language	18%	27%	32%	23%	
Staff use of books with children	7%	34%	30%	30%	
Encouraging children's use of books	7%	9%	72%	12%	
Becoming familiar with print	2%	2%	39%	57%	
Learning Activities					
Fine motor	9%	55%	23%	14%	
Art	0%	23%	52%	25%	
Music and movement	0%	16%	39%	45%	
Blocks	0%	18%	39%	43%	
Dramatic play	0%	25%	36%	39%	
Nature/science	0%	7%	55%	39%	
Math materials and activities	0%	0%	41%	59%	
Math in daily events	0%	2%	39%	59%	
Understanding written numbers	2%	0%	5%	93%	
Promoting acceptance of diversity	0%	45%	50%	5%	
Appropriate use of technology	0%	2%	18%	32%	48%
Interaction					
Supervision of gross motor	23%	39%	25%	14%	
Individualized teaching and learning	27%	32%	27%	14%	
Staff-child interaction	39%	11%	32%	18%	
Peer interactions	20%	36%	23%	20%	
Discipline	25%	27%	20%	27%	
Program Structure					
Transitions	34%	16%	25%	25%	
Free play	14%	23%	57%	7%	
Whole-group activities for play and learning	19%	30%	28%	23%	

For several items, there are one or two specific indicators driving low final item scores on the ITERS and ECERS for many of the CDC classrooms. Items for which over 25% of the classrooms scored Less than Minimal and for which 20% or more of the classrooms had a poor score on a particular indicator for the ITERS and ECERS are listed in in Table 36 and Table 37, respectively.

Table 36
Indicators Driving Low Scores on ITERS Items

Subscale Item (% of classrooms that scored less than minimal on the item) Indicator	% of classrooms with a poor score on the indicator
Space and Furnishing	
Display for children (40%)	
1.3 Staff do not talk to children about displayed materials during the observation.	36%
Personal Care Routines	
Health Practices (74%)	
1.3 Little attempt to ensure that nap/rest provisions are sanitary.	64%
3.3 Nap/rest provisions are sanitary; only minor problems; most cribs/cots are 18" apart.	61%
Language and Books	
Encouraging children's use of books (27%)	
3.3 Staff show some positive, and no negative, interest when child uses books independently.	20%
Activities	
Art (31%)	
3.4 Staff name colors as children use art materials.	20%
Blocks (51%)	
1.2 Staff show no interest in children's block play or structures.	40%
3.3 Some positive involvement by staff when children use blocks.	49%
Dramatic Play (53%)	
3.4 Staff name some objects children experience in their dramatic play.	52%
Nature/science (49%)	
3.2 Some opportunity to experience the natural world or natural objects (outdoors or indoors).	21%
3.3 Some supervision when children use nature/science materials, including sand/water if provided.	42%
Math/number (57%)	
3.2 Staff sometimes talk about shape or size when children use materials.	47%
3.3 Staff sometimes point to each item as they count for children.	36%

Table 37

Indicators Driving Low Scores on ECERS Items

Subscale Item (% of classrooms that scored less than minimal on the item) Indicator	% of classrooms with a poor score on the indicator
Space and Furnishing	
Child-related display (50%)	
1.3 Staff do not talk about display with the children.	34%
3.3 Staff talk about display materials at least once during the observation.	48%
Personal Care Routines	
Health practices (32%)	
1.2 Little attempt to ensure that nap/rest provisions are sanitary.	25%
3.2 Some attempt made to practice sanitary nap procedures.	30%
Language and Literacy	
Staff use of books with children (30%)	
3.3 The majority of children appear to be engaged for most of the time when books are used.	20%
Becoming familiar with print (57%)	
3.2 Staff point out and read print to children.	52%
Learning Activities	
Music and movement (45%)	
3.1 At least 3 music materials accessible to the children for at least 25 minutes during the observation.	34%
Blocks (43%)	
1.2 Staff show little or no interest in children's block play.	25%
3.4 Some positive involvement by staff when children use blocks.	41%
Dramatic play (39%)	
1.3 Staff usually ignore children in the dramatic play area, except to stop disruptive behavior.	20%
3.2 Staff are somewhat responsive to the children during dramatic play.	36%
3.3 Most of the staff interaction is positive or neutral.	36%
Nature/science (39%)	
3.3 Sand or water, with appropriate toys, is accessible for at least 25 minutes during the observation.	32%
Math materials and activities (59%)	
1.2 Staff are never observed to show children how to use math materials, or participate when materials are used in play.	23%
3.2 Staff sometimes give information or ask basic questions about math as children play with the materials.	41%
3.3 Math activities used engage most of the participating children.	43%
Math in daily events (59%)	
3.2 Staff sometimes use math talk as children play with non-math materials in non-math areas.	36%
3.3 Staff use math talk referring to daily events during large-group time.	48%
Understanding written numbers (93%)	
3.1 Some print numbers in display materials are accompanied by pictures that show what the number means.	23%
3.3 When children play with materials credited in 3.2, staff sometimes point out the numbers and talk about them in a way that interests children.	84%
3.4 Staff sometimes relate print numbers to corresponding number of pictures or objects.	75%
Appropriate use of technology (32%)	
1.3 No staff involvement during the observation in use of electronic media beyond starting the equipment.	27%

Table 38 shows the mean (i.e., average) scores on each ITERS item for the CDCs and the civilian centers. This table also illustrates items for which there is a significant difference between the CDCs and the civilian centers and the effect size. Statistical significance tests indicate whether the results likely happened by chance. If something is statistically significant ($p < .05$), then it likely did not happen by chance (i.e., less than 5%). Effect sizes show the magnitude of the difference. That is, whether the difference between the CDC and the civilian center, for a particular item, is large or small. An effect size of .2 is considered *small*; an effect size of .5 is *medium*; an effect size of .8 is *large* (Cohen, 1988). In Table 38, statistically significant differences are in blue (better) and red (worse).

In Table 38, scores on 29 of the 33 items are higher for the CDCs than for the civilian centers. For six of these items, the difference is statistically significant, and effect sizes are medium or large. These six items are as follows:

- Indoor space
- Encouraging children's use of books
- Fine motor
- Gross motor
- Supervision of gross motor play
- Supervision of play and learning (non-gross motor)

Moreover, independent of statistical significance, many of the effect sizes are considered medium or large. For 10 items, the effect size is medium; for 5 items, the effect size is large.

Similarly, Table 39 displays the mean (i.e., average) scores on each ECERS item for the CDCs and the civilian centers. This table also indicates the items for which there is a statistically significant difference between the CDCs and the civilian centers and the effect size. In Table 39, scores on 34 of the 35 items are higher for the CDCs than for the civilian centers. For 13 of these items, the difference is statistically significant, and all but 1 have a large effect size. These 13 items are as follows:

- Indoor space
- Furnishings for care, play, and learning
- Room arrangement for play and learning
- Space for gross motor play
- Gross motor equipment
- Meals/snacks
- Toilet/diapering
- Health practices
- Safety practices
- Blocks
- Nature/science
- Promoting acceptance of diversity
- Supervision of gross motor play

Moreover, independent of statistical significance, many of the effect sizes are considered medium or large. For 8 items, the effect size is medium; for 13 items, the effect size is large.

Table 38

Differences between CDCs and Civilian Centers on ITERS Item Score Means

Items	CDC (n=98) Mean (SD)	Civilian (n=15) Mean (SD)	*p*	Effect Size
Space and Furnishing				
Indoor space	5.31 (1.75)	3.53 (1.96)	0.024	1.00
Furnishings for care, play, and learning	5.42 (1.73)	4.67 (1.72)	0.302	0.44
Room arrangement	5.34 (1.72)	5.13 (2.00)	0.723	0.12
Display for children	3.27 (1.95)	2.53 (2.03)	0.171	0.37
Personal Care Routines				
Meals/snacks	4.52 (1.66)	2.93 (1.91)	0.059	0.94
Diapering/toileting	4.32 (1.26)	3.53 (1.19)	0.065	0.62
Health practices	2.00 (1.74)	2.07 (1.49)	0.905	-0.04
Safety practices	5.74 (1.49)	4.64 (1.45)	0.065	0.74
Language and Books				
Talking with children	5.64 (1.76)	4.47 (2.17)	0.059	0.65
Encouraging vocabulary development	4.02 (1.04)	3.27 (1.16)	0.065	0.72
Responding to children's communication	5.05 (1.74)	4.07 (2.12)	0.141	0.55
Encouraging children to communicate	5.40 (1.51)	4.87 (2.10)	0.341	0.33
Staff use of books with children	4.17 (2.00)	3.33 (1.91)	0.081	0.42
Encouraging children's use of books	3.85 (1.60)	2.60 (1.45)	0.002	0.79
Activities				
Fine motor	4.28 (1.63)	3.13 (1.77)	0.024	0.69
Art	3.30 (1.74)	4.33 (2.12)	0.171	-0.58
Music and movement	3.66 (1.47)	3.73 (1.49)	0.905	-0.05
Blocks	2.91 (1.92)	2.33 (1.95)	0.257	0.30
Dramatic play	3.15 (1.68)	2.33 (1.84)	0.257	0.48
Nature/science	2.92 (1.58)	1.87 (1.64)	0.055	0.66
Math/number	2.62 (1.47)	2.47 (1.55)	0.824	0.11
Appropriate use of technology	2.67 (1.51)	N/A	-	-
Promoting acceptance and diversity	3.61 (1.52)	3.00 (1.60)	0.103	0.40
Gross motor	5.31 (1.85)	3.80 (2.08)	0.002	0.80
Interaction				
Supervision of gross motor play	5.49 (1.61)	3.60 (2.06)	0.024	1.13
Supervision of play and learning (non-gross motor)	5.15 (1.85)	3.80 (2.01)	0.038	0.73
Peer interaction	4.71 (1.31)	4.33 (1.72)	0.404	0.28
Staff-child interaction	5.47 (1.81)	4.67 (2.38)	0.262	0.42
Providing physical warmth/touch	5.46 (1.79)	4.20 (1.93)	0.055	0.70
Guiding children's behavior	4.28 (1.71)	3.13 (2.26)	0.065	0.64
Program Structure				
Schedule and transitions	4.44 (2.27)	2.36 (2.06)	0.055	0.92
Free play	5.18 (1.44)	4.60 (1.80)	0.347	0.38
Group play activities	5.12 (2.01)	4.57 (2.70)	0.695	0.26

Note. For the tests of statistical significance, generalized estimating equations were used to account for the nested nature of the data, and the false discovery rate was used to correct for multiple comparisons. Red font and blue font indicate statistically significant differences ($p < .05$). Blue indicates better scores; red indicates worse scores. Higher scores indicate higher quality.

Table 39

Differences between CDCs and Civilian Centers on ECERS Item Score Means

Items	CDC (n=44) Mean (SD)	Civilian (n=13) Mean (SD)	*p*	Effect Size
Space and Furnishing				
Indoor space	5.75 (1.88)	3.62 (1.71)	0.009	1.16
Furnishings for care, play, and learning	5.11 (1.56)	3.62 (1.50)	0.022	0.97
Room arrangement for play and learning	5.70 (1.41)	3.46 (1.81)	0.002	1.49
Space for privacy	5.39 (1.60)	4.08 (1.75)	0.057	0.80
Child-related display	2.84 (1.72)	2.54 (1.61)	0.661	0.18
Space for gross motor play	5.75 (1.66)	2.77 (1.64)	0.000	1.80
Gross motor equipment	4.89 (1.91)	1.92 (1.61)	0.000	1.61
Personal Care Routines				
Meals/snacks	4.89 (1.47)	2.85 (1.07)	0.000	1.47
Toileting/diapering	4.36 (1.69)	2.85 (1.57)	0.002	0.91
Health practices	3.48 (1.82)	1.77 (1.79)	0.034	0.94
Safety practices	6.00 (1.40)	3.67 (1.92)	0.002	1.54
Language and Literacy				
Helping children expand vocabulary	3.70 (1.58)	3.54 (1.61)	0.772	0.10
Encouraging children to use language	4.36 (1.99)	3.69 (1.60)	0.340	0.35
Staff use of books with children	3.93 (1.97)	3.85 (2.08)	0.924	0.04
Encouraging children's use of books	4.05 (1.23)	3.85 (1.86)	0.821	0.14
Becoming familiar with print	2.61 (1.26)	3.23 (1.36)	0.356	-0.48
Learning Activities				
Fine motor	4.75 (1.69)	3.69 (1.89)	0.196	0.61
Art	3.64 (1.40)	3.31 (1.75)	0.664	0.22
Music and movement	3.25 (1.31)	2.92 (1.71)	0.661	0.23
Blocks	3.07 (1.59)	1.92 (1.19)	0.020	0.76
Dramatic play	3.27 (1.96)	2.54 (1.45)	0.292	0.40
Nature/science	2.77 (1.29)	1.77 (0.83)	0.015	0.83
Math materials and activities	2.16 (1.18)	1.46 (0.97)	0.153	0.61
Math in daily events	2.32 (1.14)	2.31 (1.18)	0.977	0.01
Understanding written numbers	2.02 (0.98)	1.85 (0.90)	0.661	0.18
Promoting acceptance of diversity	4.23 (1.05)	3.38 (0.96)	0.043	0.82
Appropriate use of technology	2.22 (1.51)	1.33 (0.50)	0.065	0.61
Interaction				
Supervision of gross motor	4.93 (1.85)	2.92 (1.93)	0.020	1.08
Individualized teaching and learning	4.86 (1.89)	3.54 (1.90)	0.077	0.70
Staff-child interaction	4.77 (2.14)	3.69 (2.21)	0.194	0.50
Peer interactions	4.61 (1.99)	3.85 (1.99)	0.311	0.39
Discipline	4.43 (2.16)	3.77 (1.96)	0.340	0.31
Program Structure				
Transitions	4.61 (2.24)	3.46 (2.11)	0.081	0.52
Free play	4.57 (1.45)	3.85 (1.28)	0.153	0.51
Whole-group activities for play and learning	4.33 (2.09)	3.62 (1.71)	0.389	0.35

Note. For the tests of statistical significance, generalized estimating equations were used to account for the nested nature of the data, and the false discovery rate was used to correct for multiple comparisons. Red and blue font indicate statistically significant differences ($p < .05$). Blue indicates better scores; red indicates worse scores. Higher scores indicate higher quality.

When examining the mean differences for the ITERS subscales, the CDCs had higher scores on all six of the subscales (see Table 40). For all of the subscales, the differences were statistically significant, and effect sizes were large or medium. Furthermore, when combining all ITERS items into one total score, the average score was higher for the CDCs than for the civilian centers; this difference was statistically significant, and the effect size was large.

Table 40

Differences between CDCs and Civilian Centers on ITERS Subscale and Total Scale Averages

Scale	CDC (n=98) Mean (SD)	Civilian (n=15) Mean (SD)	*p*	Effect Size
Space and Furnishings	4.83 (1.08)	3.97 (1.18)	0.041	0.79
Personal Care Routines	4.14 (1.05)	3.24 (1.19)	0.048	0.84
Language and Books	4.69 (1.05)	3.77 (1.31)	0.024	0.85
Activities	3.52 (0.83)	2.94 (0.94)	0.045	0.69
Interaction	5.09 (1.27)	3.96 (1.54)	0.024	0.87
Program Structure	4.98 (1.55)	3.88 (1.90)	0.047	0.69
Total Scale	4.41 (0.80)	3.54 (1.07)	0.004	1.05

Note. For the tests of statistical significance, generalized estimating equations were used to account for the nested nature of the data, and the false discovery rate was used to correct for multiple comparisons. Red font and blue font indicate statistically significant differences ($p < .05$). Blue indicates better scores; red indicates worse scores. Higher scores indicate higher quality.

When examining the mean differences for the ECERS subscales, the CDCs had higher scores on all six of the subscales (see Table 41). For three of the subscales, the differences were statistically significant, and effect sizes were large or medium. These subscales were Space and Furnishings, Personal Care Routines, and Interaction. Furthermore, when combining all ECERS items into one total score, the average score was higher for the CDCs than for the civilian centers; this difference was statistically significant, and the effect size was large.

Table 41
Differences between CDCs and Civilian Centers on ECERS Subscale and Total Scale Averages

Scale	CDC (n=44) Mean (SD)	Civilian (n=13) Mean (SD)	*p*	Effect Size
Space and Furnishings	5.06 (0.85)	3.14 (0.94)	0.000	2.20
Personal Care Routines	4.68 (1.14)	2.74 (1.34)	0.000	1.63
Language and Literacy	3.74 (1.11)	3.63 (0.88)	0.710	0.10
Learning Activities	3.10 (0.87)	2.49 (0.79)	0.089	0.72
Interaction	4.72 (1.68)	3.55 (1.49)	0.048	0.71
Program Structure	4.49 (1.66)	3.64 (1.45)	0.103	0.53
Total Scale	4.13 (0.92)	3.08 (0.84)	0.001	1.17

Note. For the tests of statistical significance, generalized estimating equations were used to account for the nested nature of the data, and the false discovery rate was used to correct for multiple comparisons. Red font and blue font indicate statistically significant differences ($p < .05$). Blue indicates better scores; red indicates worse scores. Higher scores indicate higher quality.

Factors Related to Classroom Quality

Teacher education and leadership support have been identified as important features of a high-quality early childhood education. Indeed, teacher education and a supportive work environment are both included in the NAEYC standards.

Direct-Care Staff Education

As previously discussed, there are several limitations to the direct-care staff education variable (e.g., incomplete data, small group sizes at the higher levels). Therefore, although the results are consistent with theory and prior research, results from these analyses must be interpreted with **extreme caution.**

As shown in the data summary in Table 42, for the infant and toddler classrooms, the education of the direct-care staff is related to classroom quality for the Personal Care Routines, Language and Books, Activities, and Interaction subscales and for the Total Scale. That is, as the education category increases, so does the quality score. Complete data tables are available in Appendix C.

Table 42
Direct-Care Staff Education (continuous variable) and Classroom Quality

Higher Direct-Care Staff Education	
Higher quality – P*ersonal Care Routines* subscale	Infant/Toddler
Higher quality – *Language and Books* subscale	Infant/Toddler
Higher quality – *Activities* subscale	Infant/Toddler
Higher quality – *Interaction* subscale	Infant/Toddler
Higher quality – *Total Scale*	Infant/Toddler

In an effort to reduce some of the limitations related to small sample sizes and small group sizes at the top end of the education category hierarchy, the ITERS and the ECERS classrooms were combined, and the education variable was reconfigured into a dichotomous variable: 1) the classroom has at least one direct-care staff with a bachelor's degree, and 2) the classroom does not have at least one direct-care staff with a bachelor's degree. Examining the variables in this way demonstrates that those classrooms that include at least one direct-care staff member with a bachelor's degree score higher on the Language and Books/Literacy and Activities/Learning Activities subscales and on the Total Scale of classroom quality (see Table 43).

Table 43
Direct-Care Staff Education (at least one bachelor's degree in the classroom vs. no bachelor's degree) and Classroom Quality

At Least One Direct-Care Staff with a Bachelor's Degree	
Higher quality – *Language and Books/Literacy* subscale	Infant/Toddler & Preschool/Pre-K (combined)
Higher quality – Activities/*Learning Activities* subscale	Infant/Toddler & Preschool/Pre-K (combined)
Higher quality – *Total Scale*	Infant/Toddler & Preschool/Pre-K (combined)

Leadership Support

To examine the association between leadership support and classroom quality, the three leadership support items were separated into director-related support (i.e., two items) and command-related support (i.e., 1 item). As shown in the data summary in Table 44, for the Infant/Toddler classrooms, higher staff perception of director support was related to higher scores on the Interaction and Program Structure subscales of the ITERS. Similarly, higher staff perception of command support was related to higher scores on the Space and Furnishings, Language and Books, Interaction, and Program Structure subscales and the Total Scale of the ITERS. No associations between staff perception of director support or command support and the ECERS subscales were found. However, the sample size for the Preschool/Pre-K classrooms was small, which may limit the ability to detect effects.

Table 44
Leadership Support and Classroom Quality

Director Support	
Higher quality – *Interaction* subscale	Infant/Toddler
Higher quality – *Program Structure* subscale	Infant/Toddler
Command Support	
Higher quality – *Space and Furnishings* subscale	Infant/Toddler
Higher quality – *Language and Books* subscale	Infant/Toddler
Higher quality – *Interaction* subscale	Infant/Toddler
Higher quality – *Program Structure* subscale	Infant/Toddler
Higher quality – *Total Scale*	Infant/Toddler

Child Outcomes

Differential Attrition

When examining data across time, investigating whether children who attritted were different from children who did not attrit is essential. In this evaluation, for multiple variables, the children whose parents did not complete later waves were different from children of parents who did complete later waves. This is called differential attrition. Moreover, examining differential attrition within the comparison groups is also important. For the civilian centers, those children whose parents did not complete the later waves of questionnaires were doing worse at Wave 1 than those children whose parents did complete later waves of questionnaires. This was not the case for the CDCs. In subsequent analyses of differences between groups, this can lead to the appearance that the children in the civilian centers did better over time than the children in the CDCs. However, by examining differential attrition, it is apparent that is not the case.

Table 45 illustrates that the children in the civilian centers whose parents did not complete the Wave 4 questionnaire had more emotional symptoms, conduct problems, hyperactivity, peer problems, internalizing behaviors, externalizing behaviors, and total difficulties at Wave 1. Similarly, the children in the civilian centers whose parents did not complete the Wave 5 questionnaire had more emotional symptoms, hyperactivity, and more externalizing behaviors at Wave 1. Table 45 only includes Wave 4 and Wave 5 because there was no evidence of differential attrition for those who attritted at Wave 2 or Wave 3.

Table 45
SDQ 2 & 4 (Combined) Differential Attrition

	CDCs			Civilian Centers		
	Estimated Marginal Means (Std. Error)			Estimated Marginal Means (Std. Error)		
	Attritted	Did not Attrit	*p*	Attritted	Did not Attrit	*p*
Differential Attrition at Wave 4						
W1 Emotional Symptoms	0.88 (0.20)	1.09 (0.13)	.368	3.00 (0.58)	0.59 (0.28)	.000*
W1 Conduct Problems	1.26 (0.23)	1.24 (0.15)	.925	3.20 (0.68)	1.14 (0.32)	.007*
W1 Hyperactivity	2.95 (0.34)	3.42 (0.21)	.250	5.80 (0.98)	2.46 (0.47)	.002*
W1 Peer Problems	1.20 (0.21)	1.26 (0.13)	.782	2.80 (0.60)	0.86 (0.29)	.004*
W1 Prosocial Behaviors	8.69 (0.24)	8.45 (0.15)	.397	8.60 (0.69)	8.64 (0.33)	.962
W1 Internalizing	2.07 (0.35)	2.36 (0.22)	.485	5.80 (0.99)	1.46 (0.47)	.000*
W1 Externalizing	4.24 (0.50)	4.65 (0.31)	.491	9.00 (1.43)	3.59 (0.68)	.001*
W1 Total Difficulties	6.32 (0.71)	7.01 (0.44)	.410	14.80 (2.04)	5.05 (0.97)	.000*
Differential Attrition at Wave 5						
W1 Emotional Symptoms	0.96 (0.19)	1.07 (0.13)	.626	2.57 (0.50)	0.50 (0.29)	.000*
W1 Conduct Problems	1.45 (0.22)	1.14 (0.15)	.252	2.43 (0.58)	1.20 (0.34)	.069
W1 Hyperactivity	2.96 (0.32)	3.44 (0.22)	.210	5.00 (0.83)	2.40 (0.49)	.008*
W1 Peer Problems	1.35 (0.20)	1.19 (0.14)	.502	2.00 (0.52)	0.95 (0.31)	.083
W1 Prosocial Behaviors	8.55 (0.22)	8.51 (0.16)	.864	8.71 (0.58)	8.60 (0.34)	.866
W1 Internalizing	2.31 (0.33)	2.26 (0.23)	.900	4.57 (0.85)	1.45 (0.50)	.002
W1 Externalizing	4.44 (0.47)	4.59 (0.33)	.795	7.43 (1.23)	3.60 (0.73)	.008*
W1 Total Difficulties	6.75 (0.67)	6.85 (0.47)	.904	12.00 (1.76)	5.05 (1.04)	.001

Note. Estimated marginal means are means adjusted for all other variables in the model. An * indicates that there is a statistically significant difference between families who attritted and families who did not attrit on the indicated subscale. A *p* value of < .05 was used when Levene's Test of Equality of Error Variances was not significant; a *p* value of <.001 was used when Levene's Test of Equality of Error Variances was significant. Red indicates on which subscales the attritted group has worse scores than the non-attritted group at Wave 1. For all subscales except for Prosocial Behaviors, higher scores indicate more problematic behavior. For the Prosocial Behaviors subscale, higher scores indicate more prosocial behavior.

Bivariate Correlations

Bivariate correlations are a simple measure of the relationship between two variables. Directionality cannot be inferred from correlational data. That is, if parent stress is correlated with child externalizing behaviors, there is no way to know, from this type of analysis, whether parent stress causes child externalizing behaviors, whether child externalizing behaviors causes parent stress, or if a third variable is causing both. Furthermore, these analyses include no other variables (i.e., no other variables are controlled for) and are only used as preliminary analyses to inform subsequent analyses. Bivariate correlations for the child outcomes and covariates at Wave 1 are presented in Table 46. The full correlation table, for all five waves of data, is available in Appendix C.

Table 46

Bivariate Correlations – Covariates and Child Outcome Variables Wave 1

		Child Sex	Children of Color (race)	Children of Color (ethnicity)	Parent Education	Parent Stress	Number of Relocations	Major Life Changes
TS Gold	Social-Emotional	ns	-.231**	ns	ns	ns	ns	ns
	Physical	ns	ns	ns	ns	ns	ns	ns
	Language	ns	-.257**	ns	ns	ns	ns	ns
	Cognitive	ns	-.216**	-.177*	.159*	ns	ns	ns
	Literacy	ns	ns	ns	.232**	ns	ns	ns
	Math	ns	-.237**	ns	.198**	ns	ns	ns
ASQ	Communication	ns	.178**	ns	ns	ns	ns	ns
	Gross Motor	ns	ns	ns	ns	ns	ns	ns
	Fine Motor	ns	ns	ns	ns	ns	ns	ns
	Problem Solving	ns	ns	ns	ns	ns	ns	ns
	Personal Social	-.161*	ns	ns	ns	ns	-.134*	ns
	Social Emotional	-.135*	ns	ns	ns	.250**	ns	ns
SDQ (age 2-3)	Emotional Symptoms	ns	ns	ns	ns	.238*	ns	ns
	Conduct Problems	ns	ns	.217*	ns	.363**	ns	ns
	Hyperactivity	-.237*	ns	ns	ns	ns	ns	ns
	Peer Problems	-.242**	ns	.186*	ns	ns	.283**	ns
	Prosocial Behaviors	.290**	ns	ns	ns	ns	ns	ns
	Internalizing	ns	ns	ns	ns	.223*	.226*	ns
	Externalizing	-.198*	ns	ns	ns	.276**	ns	ns
	Total Difficulties	-.217*	ns	ns	ns	.297**	.232*	ns
SDQ (age 4-5)	Emotional Symptoms	ns	ns	ns	ns	.330*	ns	ns
	Conduct Problems	-.284*	ns	ns	ns	ns	ns	ns
	Hyperactivity	ns	ns	ns	ns	ns	ns	ns
	Peer Problems	ns	ns	ns	ns	ns	ns	ns
	Prosocial Behaviors	ns	ns	ns	ns	ns	ns	.379**
	Internalizing	ns	ns	ns	ns	ns	ns	ns
	Externalizing	ns	ns	ns	ns	ns	ns	ns
	Total Difficulties	ns	ns	ns	ns	ns	ns	ns
SDQ (age 2-5)	Emotional Symptoms	ns	ns	ns	ns	.260**	ns	ns
	Conduct Problems	ns	ns	.155*	ns	.310**	ns	ns
	Hyperactivity	ns	ns	ns	ns	ns	ns	ns
	Peer Problems	ns	ns	ns	ns	ns	.196**	ns
	Prosocial Behaviors	.253**	ns	ns	ns	-.182*	ns	ns
	Internalizing	ns	ns	ns	ns	.230**	.183*	ns
	Externalizing	-.162*	ns	ns	ns	.235**	ns	ns
	Total Difficulties	-.159*	ns	ns	ns	.272**	ns	ns
EDI	Language and Cognitive	ns	ns	ns	ns	ns	ns	ns
	Physical	ns	ns	ns	.372**	-.326*	ns	ns
	Communication Skills	ns	ns	ns	ns	ns	.303*	ns

Note. *p < .05; **p < .01; ns = not significant. Child Sex: male = 0; female = 1. Children of Color (race): White = 0; Black or African American, American Indian or Alaska Native, or Asian = 1. Children of Color (ethnicity): Yes, Mexican, Mexican-American, Chicano, Puerto Rican, Cuban, or other Spanish/Hispanic/Latino = 0; No, not Spanish/Hispanic/Latino = 1. TS Gold: higher scores indicate more advanced levels of development. ASQ-3 and ASQ:SE: higher scores indicate potential delayed development. SDQ: for the Prosocial Behaviors subscale, higher scores indicate more prosocial behavior; for all other subscales, higher scores indicate more problematic behavior. EDI: higher scores indicate more advanced levels of development.

Child Development – Direct-Care Staff Report Descriptive Statistics and Norms

Initial analyses of the direct-care staff-reported child development data show the percent of children in each child development category. Data are separated by age group, domain of development, and wave of data collection (see Tables 47, 48, and 49).

Table 47

Percent of Children in Each TS Gold Category – Wave 1

	CDCs			
	Progressing Towards Expectations	Meeting Expectations	Exceeding Expectations	n
Wave 1				
Birth to 1 Year Old				
Social-Emotional	6%	76%	18%	33
Physical	24%	67%	9%	33
Language	29%	71%	0%	31
Cognitive	13%	77%	10%	30
Literacy	0%	97%	3%	30
Math	0%	97%	3%	31
1 to 2 Years Old				
Social-Emotional	26%	70%	4%	46
Physical	43%	49%	9%	47
Language	47%	51%	2%	47
Cognitive	20%	78%	2%	46
Literacy	38%	42%	20%	45
Math	61%	33%	7%	46
2 to 3 Years Old				
Social-Emotional	23%	69%	8%	48
Physical	12%	80%	8%	50
Language	23%	70%	6%	47
Cognitive	9%	82%	9%	44
Literacy	42%	49%	10%	41
Math	40%	53%	8%	40
Preschool 3				
Social-Emotional	32%	51%	17%	90
Physical	27%	57%	17%	90
Language	28%	60%	12%	89
Cognitive	36%	54%	10%	89
Literacy	42%	53%	6%	89
Math	38%	55%	7%	89
Pre-K 4				
Social-Emotional	44%	31%	25%	16
Physical	44%	44%	13%	16
Language	44%	50%	6%	16
Cognitive	38%	56%	6%	16
Literacy	44%	50%	6%	16
Math	50%	44%	6%	16

Note. Due to rounding, totals may not equal 100%.

Table 48

Percent of Children in Each TS Gold Category – Wave 2

	CDCs			
	Progressing Towards Expectations	Meeting Expectations	Exceeding Expectations	n
Wave 2				
Birth to 1 Year Old				
Social-Emotional	0%	88%	13%	24
Physical	0%	100%	0%	24
Language	8%	92%	0%	24
Cognitive	0%	78%	22%	23
Literacy	0%	83%	17%	23
Math	0%	83%	17%	23
1 to 2 Years Old				
Social-Emotional	22%	75%	3%	36
Physical	30%	65%	5%	37
Language	47%	53%	0%	36
Cognitive	20%	77%	3%	35
Literacy	34%	54%	11%	35
Math	69%	29%	3%	35
2 to 3 Years Old				
Social-Emotional	21%	73%	6%	48
Physical	19%	75%	6%	48
Language	35%	60%	4%	48
Cognitive	20%	74%	7%	46
Literacy	47%	44%	9%	45
Math	40%	53%	7%	45
Preschool 3				
Social-Emotional	17%	60%	23%	60
Physical	20%	59%	21%	61
Language	23%	58%	18%	60
Cognitive	14%	73%	14%	59
Literacy	29%	59%	13%	56
Math	20%	61%	19%	54
Pre-K 4				
Social-Emotional	35%	42%	23%	43
Physical	37%	42%	21%	43
Language	37%	51%	12%	43
Cognitive	47%	42%	12%	43
Literacy	35%	54%	12%	43
Math	44%	49%	7%	43

Note. Due to rounding, totals may not equal 100%.

Table 49

Percent of Children in Each TS Gold Category – Wave 3

	CDCs			
	Progressing Towards Expectations	Meeting Expectations	Exceeding Expectations	n
Wave 3				
Birth to 1 Year Old				
Social-Emotional	0%	50%	50%	6
Physical	0%	100%	0%	6
Language	0%	83%	17%	6
Cognitive	0%	33%	67%	6
Literacy	0%	100%	0%	5
Math	0%	83%	17%	6
1 to 2 Years Old				
Social-Emotional	18%	82%	0%	22
Physical	32%	68%	0%	22
Language	32%	68%	0%	22
Cognitive	9%	91%	0%	22
Literacy	23%	68%	9%	22
Math	59%	41%	0%	22
2 to 3 Years Old				
Social-Emotional	26%	67%	7%	27
Physical	19%	74%	7%	27
Language	30%	63%	7%	27
Cognitive	19%	82%	0%	27
Literacy	42%	42%	15%	26
Math	39%	58%	4%	26
Preschool 3				
Social-Emotional	23%	67%	10%	48
Physical	21%	57%	21%	47
Language	19%	68%	13%	47
Cognitive	28%	64%	9%	47
Literacy	43%	50%	7%	44
Math	42%	44%	13%	45
Pre-K 4				
Social-Emotional	46%	42%	13%	24
Physical	42%	54%	4%	24
Language	33%	63%	4%	24
Cognitive	50%	46%	4%	24
Literacy	46%	54%	0%	24
Math	42%	54%	4%	24

Note. Due to rounding, totals may not equal 100%.

The next set of analyses compares the children's TS Gold scores to national norms. The normed data vary based on the checkpoint; therefore, for these analyses, the children are grouped by checkpoint instead of wave of data collection. TS Gold does not provide normed data for the Summer checkpoint, only Fall, Winter, and Spring checkpoints. As such, CDC data are presented without a comparison in Table 50. The comparison of participant scores to normed data for the Fall 2019/2020 checkpoint is shown in Table 51, and the Winter 2019/2020 checkpoint is illustrated in Table 52. These tables indicate that, for some outcomes and for some age ranges, the children in the CDCs are doing better than the normed sample. However, for other outcomes and age ranges, the normed sample is doing better. All differences that are shown with blue font indicate better than the normed sample, and red font indicates worse than the normed sample. Please note, some differences are quite small, while other differences are large. The Meeting Expectations category is not easily compared. Therefore, only the Progressing Towards Expectations and the Exceeding Expectations categories are compared to the normed data.

For the Fall checkpoint, children in the Birth to 1, Preschool 3, and Pre-K 4 classrooms are doing better than the normed sample in most domains. For the 1- to 2-year-old and 2- to 3-year-old classrooms, the scores on several domains are split in two different ways: (1) there are fewer children Progressing Towards Expectations than in the normed sample, but there are also fewer children Exceeding Expectations, or (2) there are more children Exceeding Expectations than in the normed sample, but there are also more children Progressing Towards Expectations.

For the Winter checkpoint, children in the Birth to 1-year-old classrooms were doing better than the normed sample on four of the six domains. For the 1- to-2-year-old and the 2- to 3-year-old classrooms, children were generally doing worse than the normed sample. In the Preschool 3 and Pre-K 4 classrooms, the results were split. That is, there were more children Exceeding Expectations than in the normed sample, but there were also more children Progressing Towards Expectations.

Table 50

TS Gold Compared to the Normed Sample - Summer 2018/2019 Checkpoint (Cohort 1 Wave 1)

	CDCs				Normed Sample		
	Progressing Towards Expectations	Meeting Expectations	Exceeding Expectations	n	Progressing Towards Expectations	Meeting Expectations	Exceeding Expectations
Summer 2018/2019							
Birth to 1 Year Old							
Social-Emotional	9%	77%	14%	22	n/a	n/a	n/a
Physical	27%	68%	5%	22	n/a	n/a	n/a
Language	23%	77%	0%	22	n/a	n/a	n/a
Cognitive	10%	76%	14%	21	n/a	n/a	n/a
Literacy	0%	100%	0%	22	n/a	n/a	n/a
Math	0%	100%	0%	22	n/a	n/a	n/a
1 to 2 Years							
Social-Emotional	23%	70%	7%	30	n/a	n/a	n/a
Physical	42%	48%	10%	31	n/a	n/a	n/a
Language	45%	52%	3%	31	n/a	n/a	n/a
Cognitive	17%	80%	3%	30	n/a	n/a	n/a
Literacy	38%	45%	17%	29	n/a	n/a	n/a
Math	60%	33%	7%	30	n/a	n/a	n/a
2 to 3 Years							
Social-Emotional	27%	70%	3%	30	n/a	n/a	n/a
Physical	16%	78%	6%	32	n/a	n/a	n/a
Language	31%	62%	7%	29	n/a	n/a	n/a
Cognitive	14%	71%	14%	28	n/a	n/a	n/a
Literacy	41%	44%	15%	27	n/a	n/a	n/a
Math	46%	42%	12%	26	n/a	n/a	n/a
Preschool 3							
Social-Emotional	36%	48%	16%	77	n/a	n/a	n/a
Physical	29%	60%	12%	77	n/a	n/a	n/a
Language	31%	60%	9%	75	n/a	n/a	n/a
Cognitive	38%	54%	8%	76	n/a	n/a	n/a
Literacy	45%	50%	5%	76	n/a	n/a	n/a
Math	39%	54%	7%	76	n/a	n/a	n/a
Pre-K 4							
Social-Emotional	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Physical	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Language	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Cognitive	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Literacy	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Math	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Note. Due to rounding, totals may not equal 100%. Normed data are not available from TS Gold for the Summer checkpoint.

Table 51

TS Gold Compared to the Normed Sample - Fall 2019/2020 Checkpoint (Cohort 1 Wave 2, Cohort 2 Wave 1)

	CDCs				Normed Sample		
	Progressing Towards Expectations	Meeting Expectations	Exceeding Expectations	n	Progressing Towards Expectations	Meeting Expectations	Exceeding Expectations
Fall 2019/2020							
Birth to 1 Year Old							
Social-Emotional	0%	80%	20%	25	9%	86%	5%
Physical	8%	84%	8%	25	21%	75%	4%
Language	26%	74%	0%	23	22%	77%	2%
Cognitive	9%	73%	18%	22	13%	83%	3%
Literacy	0%	86%	14%	21	0%	93%	7%
Math	0%	82%	18%	22	0%	97%	3%
1 to 2 Years							
Social-Emotional	25%	73%	3%	40	30%	67%	3%
Physical	38%	58%	5%	40	32%	60%	8%
Language	43%	58%	0%	40	50%	49%	1%
Cognitive	23%	77%	0%	39	18%	79%	3%
Literacy	33%	54%	13%	39	19%	77%	4%
Math	64%	33%	3%	39	43%	55%	2%
2 to 3 Years							
Social-Emotional	22%	70%	8%	50	34%	60%	7%
Physical	16%	78%	6%	50	28%	63%	9%
Language	28%	68%	4%	50	43%	51%	6%
Cognitive	15%	81%	4%	47	32%	62%	7%
Literacy	45%	52%	2%	44	40%	53%	6%
Math	34%	64%	2%	44	40%	56%	5%
Preschool 3							
Social-Emotional	16%	64%	20%	56	40%	54%	6%
Physical	18%	56%	26%	57	31%	63%	6%
Language	21%	61%	18%	57	46%	50%	4%
Cognitive	15%	71%	15%	55	46%	50%	4%
Literacy	29%	64%	7%	56	54%	42%	4%
Math	24%	67%	9%	54	46%	50%	5%
Pre-K 4							
Social-Emotional	44%	40%	16%	43	55%	43%	2%
Physical	49%	42%	9%	43	44%	55%	1%
Language	47%	51%	2%	43	50%	49%	2%
Cognitive	47%	49%	5%	43	54%	45%	1%
Literacy	37%	60%	2%	43	53%	47%	1%
Math	49%	47%	5%	43	70%	30%	1%

Note. Blue indicates percentages that are better than the normed sample; red indicates percentages that are worse than the normed sample. Red and blue designations are only assigned for Progressing Towards Expectations and Exceeding Expectations. Designations of better or worse were determined based on visual inspection and not statistical analysis. Due to rounding, totals may not equal 100%.

Table 52

TS Gold Compared to the Normed Sample - Winter 2019/2020 Checkpoint (Cohort 1 Wave 3, Cohort 2 Wave 2)

	CDCs				Normed Sample		
	Progressing Towards Expectations	Meeting Expectations	Exceeding Expectations	n	Progressing Towards Expectations	Meeting Expectations	Exceeding Expectations
Winter 2019/2020							
Birth to 1 Year Old							
Social-Emotional	0%	75%	25%	16	2%	84%	14%
Physical	0%	100%	0%	16	4%	82%	14%
Language	0%	94%	6%	16	7%	88%	5%
Cognitive	0%	69%	31%	16	3%	86%	11%
Literacy	0%	87%	13%	15	0%	80%	20%
Math	0%	88%	13%	16	0%	93%	7%
1 to 2 Years							
Social-Emotional	21%	79%	0%	34	12%	80%	8%
Physical	29%	69%	3%	35	13%	71%	16%
Language	44%	56%	0%	34	27%	69%	4%
Cognitive	12%	85%	3%	34	5%	86%	9%
Literacy	29%	56%	15%	34	6%	83%	10%
Math	65%	32%	3%	34	22%	73%	6%
2 to 3 Years							
Social-Emotional	21%	70%	9%	43	18%	69%	13%
Physical	16%	74%	9%	43	16%	67%	17%
Language	30%	63%	7%	43	27%	61%	12%
Cognitive	17%	81%	2%	42	17%	70%	13%
Literacy	44%	39%	17%	41	25%	63%	12%
Math	41%	51%	7%	41	24%	66%	10%
Preschool 3							
Social-Emotional	20%	63%	17%	65	18%	66%	16%
Physical	22%	56%	22%	64	14%	72%	15%
Language	20%	63%	17%	64	26%	63%	11%
Cognitive	25%	64%	11%	64	23%	65%	13%
Literacy	39%	49%	12%	57	26%	63%	11%
Math	36%	43%	21%	58	21%	65%	14%
Pre-K 4							
Social-Emotional	33%	41%	26%	39	25%	66%	10%
Physical	28%	51%	21%	39	19%	75%	6%
Language	26%	59%	15%	39	24%	70%	6%
Cognitive	44%	44%	13%	39	25%	71%	4%
Literacy	41%	46%	13%	39	21%	75%	4%
Math	38%	54%	8%	39	36%	60%	4%

Note. Blue indicates percentages that are better than the normed sample; red indicates percentages that are worse than the normed sample. Red and blue designations are only assigned for Progressing Towards Expectations and Exceeding Expectations. Designations of better or worse were determined based on visual inspection and not statistical analysis. Due to rounding, totals may not equal 100%.

Child Development – Parent Report Descriptive Statistics

Table 53 shows the percent of children in each developmental category for parent reports of child development. In initial analyses comparing children from the CDCs and the civilian centers, with no covariates included, there were no statistically significant differences between the CDCs and the civilian centers regarding the percent of children in each ASQ-3 category. This measure is primarily used as a screener; therefore, as would be expected, the vast majority of children fall into the Typical Development category. One domain to note, however, is the Fine Motor domain at Wave 1. Nearly 25% of children in both the CDCs and civilian centers scored in the Monitor Zone or Referral categories at Wave 1.

Table 53

Percent of Children in Each ASQ-3 Category

	CDCs				Civilian Centers			
	Typical Development	Monitor Zone	Referral	n	Typical Development	Monitor Zone	Referral	n
Wave 1								
Communication	89%	9%	2%	214	89%	11%	0%	36
Gross Motor	85%	7%	8%	213	81%	11%	8%	36
Fine Motor	77%	13%	10%	210	77%	17%	6%	35
Problem Solving	91%	6%	4%	211	86%	11%	3%	35
Personal Social	88%	10%	2%	209	83%	9%	9%	35
Wave 2								
Communication	90%	8%	2%	169	93%	7%	0%	30
Gross Motor	87%	9%	4%	170	83%	17%	0%	30
Fine Motor	87%	10%	3%	166	93%	4%	4%	28
Problem Solving	94%	3%	3%	161	96%	4%	0%	28
Personal Social	91%	8%	1%	167	93%	7%	0%	29
Wave 3 (Cohort 2 during COVID)								
Communication	95%	5%	1%	154	94%	7%	0%	31
Gross Motor	92%	6%	3%	155	90%	7%	3%	31
Fine Motor	86%	11%	3%	154	81%	16%	3%	31
Problem Solving	94%	5%	1%	151	94%	7%	0%	31
Personal Social	96%	3%	1%	154	90%	0%	7%	31
Wave 4 (during COVID)								
Communication	94%	6%	0%	144	93%	7%	0%	28
Gross Motor	92%	4%	4%	144	93%	4%	4%	27
Fine Motor	85%	11%	4%	141	93%	7%	0%	28
Problem Solving	95%	4%	1%	141	96%	4%	0%	27
Personal Social	93%	5%	2%	143	96%	0%	4%	28
Wave 5 (during COVID)								
Communication	93%	6%	2%	133	92%	4%	4%	25
Gross Motor	93%	6%	2%	133	92%	8%	0%	25
Fine Motor	89%	7%	5%	133	92%	4%	4%	25
Problem Solving	95%	4%	2%	132	88%	8%	4%	25
Personal Social	95%	3%	2%	133	92%	4%	4%	25

Note. There were no differences between the participants in the CDCs and participants in civilian centers for this analysis. Due to rounding, totals may not equal 100%.

Table 54 presents the percent of participants in each developmental category for the parent report of social-emotional development. In initial analyses that compared participants in the CDCs and the civilian centers, with no covariates included, only one difference emerged between CDCs and civilian centers for the measure of parent-reported social-emotional skills. At Wave 2, a higher percent of children in the CDCs were in the Monitor category, and a lower percent of children were in the Low or No Risk category than were in those categories for the civilian centers.

Table 54

Percent of Children in Each ASQ:SE Category

	CDCs				Civilian Centers			
	Low or No Risk	Monitor	Refer	n	Low or No Risk	Monitor	Refer	n
Wave 1	85%	9%	6%	196	77%	12%	12%	34
Wave 2	82%	14%	4%	163	100%	0%	0%	28
Wave 3	83%	11%	7%	153	90%	0%	10%	31
Wave 4	85%	7%	7%	149	97%	3%	0%	30
Wave 5	90%	6%	3%	146	93%	0%	7%	29

Note. Red font and blue font indicate that there is a statistically significant difference ($p < .05$) between participants in the CDCs and participants in the civilian centers at the indicated level. Blue indicates better scores; red indicates worse scores. Due to rounding, totals may not equal 100%.

Child Well-Being – Parent Report Descriptive Statistics and Norms

Tables 55, 56, and 57, show mean scores on the child well-being measure. Due to small sample sizes and because the measures are nearly identical for the two age groups, for these initial analyses, the evaluation team examined the results for the SDQ split by age group and combined. In initial analyses that compared participants in the CDCs and the civilian centers on parent-reported child well-being, with no covariates included, no differences between children in CDCs and children in civilian centers emerged for child well-being at Waves 1 and 3. Differences did emerge at Waves 4 and 5. However, as discussed above, there was evidence of differential attrition. Therefore, what appears to be the children in civilian centers doing better over time, could actually be explained by the fact that the families in the civilian centers who dropped out of the evaluation had worse scores on the well-being domains at Wave 1 than those who stayed in the evaluation. There was no difference in attrition for the CDCs. Significant differences are indicated in the table with an [x]; however, the significant differences are not meaningful due to the differential attrition.

Table 55
SDQ 2 Means

	CDCs		Civilian Centers	
	Mean (SD)	n	Mean (SD)	n
Wave 1				
Emotional Symptoms	1.07 (1.29)	97	0.95 (1.67)	20
Conduct Problems	1.46 (1.56)	98	1.65 (2.32)	20
Hyperactivity	3.18 (2.18)	97	2.95 (2.24)	20
Peer Problems	1.40 (1.43)	97	2.25 (2.83)	20
Prosocial Behaviors	8.36 (1.54)	98	8.60 (1.98)	20
Internalizing Behaviors	2.47 (2.30)	97	2.25 (2.83)	20
Externalizing Behaviors	4.65 (3.27)	97	4.60 (4.28)	20
Total Difficulties	7.12 (4.55)	97	6.85 (6.79)	20
Wave 3 (Cohort 2 during COVID)				
Emotional Symptoms	1.05 (1.34)	64	0.58 (0.67)	12
Conduct Problems	1.47 (1.41)	64	1.75 (0.36)	12
Hyperactivity	3.23 (2.17)	64	2.42 (2.11)	12
Peer Problems	1.20 (1.37)	64	0.75 (0.75)	12
Prosocial Behaviors	8.58 (1.50)	64	8.50 (1.88)	12
Internalizing Behaviors	2.25 (2.34)	64	1.33 (1.23)	12
Externalizing Behaviors	4.70 (3.00)	64	4.17 (2.62)	12
Total Difficulties	6.95 (4.74)	64	5.50 (2.81)	12
Wave 4 (during COVID)				
Emotional Symptoms	0.72 (1.06)	64	0.38 (0.51)	13
Conduct Problems	1.50 (1.60)	64	1.15 (0.90)	13
Hyperactivity	3.03 (1.76)[x]	64	1.62 (1.26)[x]	13
Peer Problems	1.09 (1.26)	64	0.62 (1.04)	13
Prosocial Behaviors	8.30 (1.74)	64	8.92 (1.50)	13
Internalizing Behaviors	1.81 (1.94)	64	1.00 (1.22)	13
Externalizing Behaviors	4.53 (2.75)[x]	64	2.77 (1.59)[x]	13
Total Difficulties	6.34 (3.99)[x]	64	3.77 (2.24)[x]	13
Wave 5 (during COVID)				
Emotional Symptoms	0.76 (0.85)[x]	68	0.27 (0.47)[x]	11
Conduct Problems	0.94 (0.99)	68	1.00 (1.10)	11
Hyperactivity	2.59 (1.75)	68	1.82 (1.33)	11
Peer Problems	1.24 (1.27)[x]	68	0.45 (0.69)[x]	11
Prosocial Behaviors	8.40 (1.62)[x]	68	9.55 (0.69)[x]	11
Internalizing Behaviors	2.00 (1.60)[x]	68	0.73 (1.01)[x]	11
Externalizing Behaviors	3.53 (2.36)	68	2.82 (2.14)	11
Total Difficulties	5.53 (3.04)[x]	68	3.55 (2.42)[x]	11

Note. An [x] indicates that there is a statistically significant difference ($p < .05$) between participants in the CDCs and participants in the civilian centers on the indicated subscale. Please note, the statistically significant difference at Waves 4 and 5 are due to differential attrition, not due to actual improvements over time for the children in civilian centers. For the Prosocial Behaviors subscale, higher scores indicate more prosocial behavior; for all other subscales, higher scores indicate more problematic behavior.

Table 56
SDQ 4 Means

	CDCs		Civilian Centers	
	Mean (SD)	n	Mean (SD)	n
Wave 1				
Emotional Symptoms	0.96 (1.21)	50	1.29 (2.14)	7
Conduct Problems	0.82 (0.98)	50	1.14 (1.35)	7
Hyperactivity	3.50 (2.26)	50	3.43 (3.05)	7
Peer Problems	0.94 (1.15)	50	1.00 (1.53)	7
Prosocial Behaviors	8.84 (1.35)	50	8.71 (0.76)	7
Internalizing Behaviors	1.90 (1.87)	50	2.29 (3.55)	7
Externalizing Behaviors	4.32 (2.84)	50	4.57 (4.04)	7
Total Difficulties	6.22 (3.83)	50	6.86 (7.38)	7
Wave 3 (Cohort 2 during COVID)				
Emotional Symptoms	1.57 (1.66)	56	1.25 (2.05)	12
Conduct Problems	1.00 (1.22)	56	1.25 (1.42)	12
Hyperactivity	3.32 (2.34)	56	2.17 (2.69)	12
Peer Problems	1.14 (1.18)	56	1.08 (1.98)	12
Prosocial Behaviors	8.79 (1.40)	56	8.67 (1.56)	12
Internalizing Behaviors	2.71 (2.50)	56	2.33 (3.87)	12
Externalizing Behaviors	4.32 (3.15)	56	3.42 (3.96)	12
Total Difficulties	7.04 (4.67)	56	5.75 (7.44)	12
Wave 4 (during COVID)				
Emotional Symptoms	1.51 (1.76)[x]	65	0.23 (0.44)[x]	13
Conduct Problems	1.11 (1.44)	65	0.54 (0.78)	13
Hyperactivity	3.58 (2.39)[x]	65	1.93 (1.61)[x]	13
Peer Problems	1.25 (1.52)[x]	65	0.38 (0.65)[x]	13
Prosocial Behaviors	8.48 (1.71)	65	9.08 (1.26)	13
Internalizing Behaviors	2.75 (2.82)[x]	65	0.62 (0.87)[x]	13
Externalizing Behaviors	4.69 (3.46)[x]	65	2.46 (2.03)[x]	13
Total Difficulties	7.45 (5.24)[x]	65	3.08 (2.18)[x]	13
Wave 5 (during COVID)				
Emotional Symptoms	1.48 (1.77)[x]	67	0.54 (0.97)[x]	13
Conduct Problems	0.93 (1.20)	67	0.69 (0.95)	13
Hyperactivity	3.06 (2.26)	67	2.23 (1.92)	13
Peer Problems	1.10 (1.37)	67	0.54 (0.88)	13
Prosocial Behaviors	9.10 (1.24)	67	9.00 (1.68)	13
Internalizing Behaviors	2.58 (2.81)[x]	67	1.08 (1.50)[x]	13
Externalizing Behaviors	3.99 (3.13)	67	2.92 (2.63)	13
Total Difficulties	6.57 (5.03)	67	4.00 (3.67)	13

Note. An [x] indicates that there is a statistically significant difference ($p < .05$) between participants in the CDCs and participants in the civilian centers on the indicated subscale. Please note, the statistically significant difference at Waves 4 and 5 are due to differential attrition, not due to actual improvements over time for the children in civilian centers. For the Prosocial Behaviors subscale, higher scores indicate more prosocial behavior; for all other subscales, higher scores indicate more problematic behavior.

Table 57
SDQ 2 & 4 (Combined) Means

	CDCs		Civilian Centers	
	Mean (SD)	n	Mean (SD)	n
Wave 1				
Emotional Symptoms	1.03 (1.26)	147	1.04 (1.76)	27
Conduct Problems	1.24 (1.42)	148	1.52 (2.10)	27
Hyperactivity	3.29 (2.21)	147	3.07 (2.42)	27
Peer Problems	1.24 (1.35)	147	1.22 (1.50)	27
Prosocial Behaviors	8.52 (1.49)	148	8.63 (1.74)	27
Internalizing Behaviors	2.28 (2.17)	147	2.26 (2.96)	27
Externalizing Behaviors	4.54 (3.12)	147	4.59 (4.14)	27
Total Difficulties	6.82 (4.33)	147	6.85 (6.80)	27
Wave 3 (Cohort 2 during COVID)				
Emotional Symptoms	1.29 (1.51)	120	0.92 (1.53)	24
Conduct Problems	1.25 (1.34)	120	1.50 (1.38)	24
Hyperactivity	3.28 (2.24)	120	2.29 (2.37)	24
Peer Problems	1.18 (1.28)	120	0.92 (1.47)	24
Prosocial Behaviors	8.68 (1.45)	120	8.58 (1.69)	24
Internalizing Behaviors	2.47 (2.41)	120	1.83 (2.85)	24
Externalizing Behaviors	4.53 (3.06)	120	3.79 (3.31)	24
Total Difficulties	6.99 (4.69)	120	5.63 (5.50)	24
Wave 4 (during COVID)				
Emotional Symptoms	1.12 (1.50)[x]	129	0.31 (0.47)[x]	26
Conduct Problems	1.30 (1.53)[x]	129	0.85 (0.88)[x]	26
Hyperactivity	3.31 (2.11)[x]	129	1.77 (1.42)[x]	26
Peer Problems	1.17 (1.39)[x]	129	0.50 (0.86)[x]	26
Prosocial Behaviors	8.39 (1.72)	129	9.00 (1.36)	26
Internalizing Behaviors	2.29 (2.46)[x]	129	0.81 (1.06)[x]	26
Externalizing Behaviors	4.61 (3.12)[x]	129	2.62 (1.79)[x]	26
Total Difficulties	6.90 (4.68)[x]	129	3.42 (2.19)[x]	26
Wave 5 (during COVID)				
Emotional Symptoms	1.12 (1.42)[x]	135	0.42 (0.78)[x]	24
Conduct Problems	0.93 (1.09)	135	0.83 (1.01)	24
Hyperactivity	2.82 (2.03)	135	2.04 (1.65)	24
Peer Problems	1.17 (1.32)[x]	135	0.50 (0.78)[x]	24
Prosocial Behaviors	8.75 (1.48)	135	9.25 (1.33)	24
Internalizing Behaviors	2.29 (2.29)[x]	135	0.92 (1.28)[x]	24
Externalizing Behaviors	3.76 (2.77)	135	2.88 (2.36)	24
Total Difficulties	6.04 (4.17)[x]	135	3.79 (3.11)[x]	24

Note. An [x] indicates that there is a statistically significant difference ($p < .05$) between participants in the CDCs and participants in the civilian centers on the indicated subscale. Please note, the statistically significant difference at Waves 4 and 5 are due to differential attrition, not due to actual improvements over time for the children in civilian centers. For the Prosocial Behaviors subscale, higher scores indicate more prosocial behavior; for all other subscales, higher scores indicate more problematic behavior.

The SDQ provides categories, which are based on the mean scores, to help interpret the data. Tables 58, 59, and 60 present the percent of participants in each category. Participants who are not in the Close to Average category are demonstrating difficulties on the subscale. In initial analyses, with no covariates included, when comparing the categories for child well-being for 2- and 3- year-old children in CDCs and the civilian centers, only two significant differences emerge. As shown in Table 58, at Wave 1, the CDCs had a lower percent of children who were categorized as Very High for Total Difficulties when compared to the civilian centers. Similarly, at Wave 3, the CDCs had a lower percent of children categorized as Low for Prosocial Behaviors when compared to the civilian centers. As illustrated in Table 59, at Wave 1, the CDCs had a higher percent of children who were categorized as Close to Average for Conduct Problems, and a lower percent of children categorized as Slightly Raised for Conduct Problems, High for Emotional Symptoms and Peer Problems, and Very High for Total Difficulties, as compared to children in the civilian centers.

Due to small sample sizes and because the measures are nearly identical for the two age groups, for these initial analyses, the evaluation team examined the results for the SDQ split by age group and combined. When examining the 2- to 5-year-olds together, two significant differences emerged. As can be seen in Table 60, at Wave 1, the CDCs had a lower percent of participants who were categorized as Very High in Total Difficulties when compared to the civilian centers. Conversely, at Wave 4, the CDCs had a lower percent of children in the Close to Average category for Hyperactivity when compared to the civilian centers.

Table 58

Percent of Children in Each SDQ 2 (i.e., Children Aged 2 & 3) Category

	CDCs					Civilian Centers				
	Close to Average	Slightly Raised (Lowered)	High (Low)	Very High (Low)	n	Close to Average	Slightly Raised (Lowered)	High (Low)	Very High (Low)	n
Wave 1										
Emotional Symptoms	86%	8%	3%	1%	97	90%	5%	0%	5%	20
Conduct Problems	90%	5%	2%	3%	98	90%	0%	0%	10%	20
Hyperactivity	89%	3%	3%	5%	97	90%	5%	0%	5%	20
Peer Problems	81%	7%	7%	4%	97	85%	10%	0%	5%	20
Prosocial Behaviors	85%	8%	5%	2%	98	90%	5%	0%	5%	20
Total Difficulties	86%	8%	6%	0%	97	90%	0%	5%	5%	20
Wave 3 (Cohort 2 during COVID)										
Emotional Symptoms	88%	5%	5%	3%	64	100%	0%	0%	0%	12
Conduct Problems	86%	11%	3%	0%	64	92%	8%	0%	0%	12
Hyperactivity	88%	3%	6%	3%	64	92%	0%	8%	0%	12
Peer Problems	77%	17%	5%	2%	64	100%	0%	0%	0%	12
Prosocial Behaviors	91%	5%	2%	3%	64	83%	0%	17%	0%	12
Total Difficulties	86%	6%	6%	2%	64	100%	0%	0%	0%	12
Wave 4 (during COVID)										
Emotional Symptoms	92%	5%	2%	2%	64	100%	0%	0%	0%	13
Conduct Problems	91%	5%	2%	3%	64	100%	0%	0%	0%	13
Hyperactivity	92%	2%	6%	0%	64	100%	0%	0%	0%	13
Peer Problems	86%	8%	5%	2%	64	82%	8%	0%	0%	13
Prosocial Behaviors	83%	6%	8%	3%	64	82%	0%	8%	0%	13
Total Difficulties	92%	5%	3%	0%	64	100%	0%	0%	0%	13
Wave 5 (during COVID)										
Emotional Symptoms	96%	4%	0%	0%	68	100%	0%	0%	0%	13
Conduct Problems	99%	2%	0%	0%	68	92%	8%	0%	0%	13
Hyperactivity	96%	4%	0%	0%	68	92%	8%	0%	0%	13
Peer Problems	84%	12%	2%	3%	68	92%	8%	0%	0%	13
Prosocial Behaviors	87%	4%	7%	2%	68	85%	0%	8%	8%	13
Total Difficulties	99%	2%	0%	0%	68	92%	8%	0%	0%	13

Note. Blue font and red font indicate that there is a statistically significant difference ($p < .05$) between participants in the CDCs and participants in the civilian centers at the indicated level. Blue indicates better scores; red indicates worse scores. Due to rounding, totals may not equal 100%.

Table 59

Percent of Children in Each SDQ 4 (i.e., Children Aged 4 & 5) Category

	CDCs					Civilian Centers				
	Close to Average	Slightly Raised (Lowered)	High (Low)	Very High (Low)	n	Close to Average	Slightly Raised (Lowered)	High (Low)	Very High (Low)	n
Wave 1										
Emotional Symptoms	94%	6%	0%	0%	50	86%	0%	14%	0%	7
Conduct Problems	94%	6%	0%	0%	50	71%	29%	0%	0%	7
Hyperactivity	82%	12%	2%	4%	50	86%	0%	0%	14%	7
Peer Problems	88%	10%	0%	2%	50	86%	0%	14%	0%	7
Prosocial Behaviors	86%	6%	4%	4%	50	100%	0%	0%	0%	7
Total Difficulties	98%	2%	0%	0%	50	86%	0%	0%	14%	7
Wave 3 (Cohort 2 during COVID)										
Emotional Symptoms	89%	5%	4%	2%	56	92%	0%	0%	8%	12
Conduct Problems	91%	2%	7%	0%	56	92%	0%	8%	0%	12
Hyperactivity	84%	7%	5%	4%	56	92%	0%	0%	8%	12
Peer Problems	89%	5%	4%	2%	56	83%	0%	8%	8%	12
Prosocial Behaviors	79%	13%	7%	2%	56	83%	8%	0%	8%	12
Total Difficulties	91%	4%	4%	2%	56	92%	0%	0%	8%	12
Wave 4 (during COVID)										
Emotional Symptoms	86%	6%	6%	2%	65	100%	0%	0%	0%	13
Conduct Problems	85%	5%	11%	0%	65	100%	0%	0%	0%	13
Hyperactivity	82%	11%	5%	3%	65	100%	0%	0%	0%	13
Peer Problems	82%	8%	3%	8%	65	100%	0%	0%	0%	13
Prosocial Behaviors	77%	8%	8%	8%	65	77%	23%	0%	0%	13
Total Difficulties	86%	6%	6%	2%	65	100%	0%	0%	0%	13
Wave 5 (during COVID)										
Emotional Symptoms	84%	8%	0%	0%	67	100%	0%	0%	0%	13
Conduct Problems	88%	8%	5%	0%	67	92%	8%	0%	0%	13
Hyperactivity	85%	9%	6%	0%	67	92%	8%	0%	0%	13
Peer Problems	90%	6%	0%	5%	67	92%	8%	0%	0%	13
Prosocial Behaviors	87%	10%	2%	2%	67	85%	0%	8%	8%	13
Total Difficulties	90%	5%	5%	2%	67	92%	8%	0%	0%	13

Note. Blue font and red font indicate that there is a statistically significant difference ($p < .05$) between participants in the CDCs and participants in the civilian centers at the indicated level. Blue indicates better scores; red indicates worse scores. Due to rounding, totals may not equal 100%.

Table 60

Percent of Children in Each SDQ 2 & 4 (i.e., Children Aged 2-5) Category

	CDCs					Civilian Centers				
	Close to Average	Slightly Raised (Lowered)	High (Low)	Very High (Low)	n	Close to Average	Slightly Raised (Lowered)	High (Low)	Very High (Low)	n
Wave 1										
Emotional Symptoms	88%	8%	2%	2%	147	89%	4%	4%	4%	27
Conduct Problems	91%	5%	1%	2%	148	85%	7%	0%	7%	27
Hyperactivity	86%	6%	3%	5%	147	89%	4%	0%	7%	27
Peer Problems	84%	8%	5%	3%	147	85%	7%	4%	4%	27
Prosocial Behaviors	85%	7%	5%	3%	148	93%	4%	0%	4%	27
Total Difficulties	90%	6%	4%	0%	147	89%	0%	4%	7%	27
Wave 3 (Cohort 2 during COVID)										
Emotional Symptoms	88%	5%	4%	3%	120	96%	0%	0%	4%	24
Conduct Problems	88%	7%	5%	0%	120	92%	4%	4%	0%	24
Hyperactivity	86%	5%	6%	3%	120	92%	0%	4%	4%	24
Peer Problems	83%	12%	4%	2%	120	92%	0%	4%	4%	24
Prosocial Behaviors	85%	8%	4%	3%	120	83%	4%	8%	4%	24
Total Difficulties	88%	5%	5%	2%	120	96%	0%	0%	4%	24
Wave 4 (during COVID)										
Emotional Symptoms	89%	5%	4%	2%	129	100%	0%	0%	0%	26
Conduct Problems	88%	5%	6%	2%	129	100%	0%	0%	0%	26
Hyperactivity	87%	6%	5%	2%	129	100%	0%	0%	0%	26
Peer Problems	84%	8%	4%	5%	129	96%	3%	0%	0%	26
Prosocial Behaviors	80%	7%	8%	5%	129	85%	11%	4%	0%	26
Total Difficulties	89%	5%	5%	1%	129	100%	0%	0%	0%	26
Wave 5 (during COVID)										
Emotional Symptoms	90%	6%	4%	0%	135	100%	0%	0%	0%	24
Conduct Problems	93%	4%	2%	0%	135	96%	4%	0%	0%	24
Hyperactivity	90%	7%	3%	0%	135	96%	4%	0%	0%	24
Peer Problems	87%	9%	1%	4%	135	96%	4%	0%	0%	24
Prosocial Behaviors	87%	7%	4%	2%	135	92%	0%	4%	4%	24
Total Difficulties	94%	3%	2%	1%	135	96%	4%	0%	0%	24

Note. Blue font and red font indicate that there is a statistically significant difference ($p < .05$) between participants in the CDCs and participants in the civilian centers at the indicated level. Blue indicates better scores; red indicates worse scores. Due to rounding, totals may not equal 100%.

National norms for the SDQ are only available for the version administered to parents of the 4- and 5-year-olds. The comparisons to the normed data indicate that for Waves 1, 3, and 5, children in the CDCs are doing better than the normed sample for several domains. At Wave 1, children in the CDCs are doing better than the normed sample on Emotional Symptoms, Conduct Problems, Peer Problems, Prosocial Behaviors, and Total Difficulties (see Table 61). At Wave 3, children in CDC are doing better than the normed sample on Conduct Problems and Prosocial

Behaviors. At Wave 5, children in the CDCs are doing better than the normed sample on Conduct Problems and Prosocial Behaviors. For the other domains, at all reported timepoints, the children in the CDCs were not different than the normed sample. For Waves 1 and 3, the children in the civilian centers were no different than the normed sample. Although there was a statistically significant difference at Waves 4 and 5, these differences are due to differential attrition, not due to actual improvements over time for the children in civilian centers.

Table 61
SDQ 4 Comparison to Normed Data

	CDCs			Civilian Centers		
	Mean Difference	*p* value	n	Mean Difference	*p* value	n
Wave 1						
Emotional Symptoms	-0.54	0.003	50	-0.21	0.800	7
Conduct Problems	-0.58	<.001	50	-0.26	0.631	7
Hyperactivity	0.30	0.353	50	0.23	0.849	7
Peer Problems	-0.36	0.032	50	-0.30	0.622	7
Prosocial Behaviors	0.44	0.025	50	0.31	0.313	7
Total Difficulties	-1.18	0.034	50	-0.54	0.852	7
Wave 3 (Cohort 2 during COVID)						
Emotional Symptoms	0.07	0.749	56	-0.25	0.681	12
Conduct Problems	-0.40	0.017	56	-0.15	0.722	12
Hyperactivity	0.12	0.700	56	-1.03	0.210	12
Peer Problems	-0.16	0.324	56	-0.22	0.711	12
Prosocial Behaviors	0.39	0.044	56	0.27	0.565	12
Total Difficulties	-0.36	0.562	56	-1.65	0.458	12
Wave 4 (during COVID)						
Emotional Symptoms	0.01	0.972	65	-1.27[x]	<.001	13
Conduct Problems	-0.29	0.106	65	-0.86[x]	0.002	13
Hyperactivity	0.38	0.199	65	-1.28[x]	0.014	13
Peer Problems	-0.05	0.776	65	-0.92[x]	<.001	13
Prosocial Behaviors	0.08	0.719	65	0.68	0.076	13
Total Difficulties	0.05	0.944	65	-4.32[x]	<.001	13
Wave 5 (during COVID)						
Emotional Symptoms	-0.02	0.918	67	-0.96[x]	0.004	13
Conduct Problems	-0.47	0.002	67	-0.71[x]	0.02	13
Hyperactivity	-0.14	0.613	67	-0.97	0.094	13
Peer Problems	-0.20	0.248	67	-0.76[x]	0.009	13
Prosocial Behaviors	0.70	<.001	67	0.60	0.223	13
Total Difficulties	-0.83	0.180	67	-3.40[x]	0.006	13

Note. Blue indicates that the CDC subsample differs statistically significantly ($p < .05$) from the normed sample. A negative number indicates that the evaluation participants have lower scores on that subscale than the normed data. A positive number indicates that the evaluation participants have higher scores on that subscale than the normed data. A [x] indicates that the results were statistically significant, but this can be explained by differential attrition rather than a true difference.

Kindergarten Readiness – Parent Report Descriptive Statistics

Table 62 presents means for parent reports of kindergarten readiness. In initial analyses comparing the CDCs and the civilian centers, with no covariates included, no differences emerged between the CDCs and the civilian centers for the measure of kindergarten readiness.

Table 62
EDI Means

	CDCs		Civilian Centers	
	Mean (SD)	n	Mean (SD)	n
Wave 1				
Language and Cognitive Development	7.02 (1.66)	48	7.48 (1.27)	7
Physical Well-Being	8.55 (1.62)	47	8.57 (1.50)	7
Communication Skills	9.01 (1.46)	47	9.05 (1.31)	7
Wave 2				
Language and Cognitive Development	7.30 (1.72)	55	7.50 (2.04)	7
Physical Well-Being	8.55 (1.76)	55	9.29 (1.31)	7
Communication Skills	9.00 (1.35)	55	9.29 (1.31)	7
Wave 3 (Cohort 2 during COVID)				
Language and Cognitive Development	7.61 (1.62)	56	7.86 (1.18)	13
Physical Well-Being	8.71 (1.75)	56	8.85 (1.42)	13
Communication Skills	9.11 (1.14)	56	9.23 (1.46)	13
Wave 4 (during COVID)				
Language and Cognitive Development	7.71 (1.67)	64	8.13 (1.49)	12
Physical Well-Being	8.97 (1.40)	65	9.03 (1.81)	12
Communication Skills	9.23 (1.32)	65	9.72 (0.65)	12
Wave 5 (during COVID)				
Language and Cognitive Development	8.24 (1.45)	67	8.49 (1.50)	13
Physical Well-Being	9.30 (1.13)	67	8.85 (1.97)	13
Communication Skills	9.42 (1.26)	67	9.49 (1.05)	13

Note. There were no differences between the participants in the CDCs and participants in civilian centers for this analysis of this measure. Higher scores indicate more advanced levels of development.

Relation Between Early Childhood Education Factors and Child Outcomes

The evaluation team examined the association between center type (i.e., CDC or civilian center), leadership support, classroom characteristics, and child outcomes. When conducting outcome analyses, the evaluation team included variables in the analyses that are not related to the child care center or the classroom quality but may affect child outcomes (i.e., covariates). The statistical model controls for the effect of these variables on the outcomes. The following covariates were included in the child outcome analyses: child sex, child race, child ethnicity, parent education, parent stress, number of relocations, and major life changes. For Waves 3 through 5 cohort was also included as a covariate. For the TS Gold subscales where there were differences between the online and the paper version, the measure version was also included as a covariate.

A summary of the Wave 1 child outcome analyses can be found in Table 63. Children in CDCs had more personal social skills, higher social emotional development, fewer conduct problems, and fewer total difficulties when compared to children in civilian centers. When CDC staff reported more director support, children had fewer conduct problems. When CDC staff reported more command support, children had higher literacy development and math development. When classroom quality was higher, children had higher physical development and language development and better problem-solving skills. Table 63 also reports for which age group the results were found. The data summary for the covariates is in Table 64.

Table 63

Wave 1 Child Outcomes

CDCs (compared to Civilian Centers)	
Higher parent reported *personal social skills*	Infant/Toddler
Higher parent reported *social emotional development*	Preschool/Pre-K
Fewer parent reported *conduct problems*	Preschool/Pre-K
Fewer parent reported *total difficulties*	Preschool/Pre-K
Director support	
Fewer parent reported *conduct problems*	Infant/Toddler and Preschool/Pre-K (combined)
Command Support	
Higher staff reported *literacy development*	Infant/Toddler and Preschool/Pre-K (combined)
Higher staff reported *math development*	Infant/Toddler and Preschool/Pre-K (combined)
Classroom Quality (Total Score)	
Higher staff reported *physical development*	Infant/Toddler
Higher staff reported *language development*	Infant/Toddler
Higher parent reported *problem-solving skills*	Infant/Toddler
Classroom Quality (Activities)	
Higher staff reported *physical development*	Infant/Toddler
Higher parent reported *problem-solving skills*	Infant/Toddler

Table 64

Wave 1 Child Covariates

Higher Parent Perceived Stress	
Lower parent reported *social emotional development*	Infant/Toddler & Preschool/Pre-K
More parent reported *conduct problems*	Infant/Toddler & Preschool/Pre-K
More parent reported *emotional symptoms*	Infant/Toddler & Preschool/Pre-K
Lower parent reported *fine motor skills*	Preschool/Pre-K
Less parent reported *prosocial behavior*	Preschool/Pre-K
More parent reported *peer problems*	Preschool/Pre-K
More parent reported *total difficulties*	Preschool/Pre-K
Higher parent reported *physical development*	Preschool/Pre-K
Parent's Higher Education	
Higher staff reported *literacy development*	Preschool/Pre-K
More Relocations During Child's Life	
Lower staff reported *language development*	Infant/Toddler
Lower staff reported *physical development*	Infant/Toddler
Higher parent reported *communication skills*	Preschool/Pre-K
More parent reported *peer problems*	Preschool/Pre-K
Child Has Special Needs or a Disability	
Lower parent reported *fine motor skills*	Infant/Toddler
Lower parent reported *gross motor skills*	Infant/Toddler
More parent reported *emotional symptoms*	Infant/Toddler
Lower parent reported *communication skills*	Infant/Toddler & Preschool/Pre-K
More parent reported *peer problems*	Preschool/Pre-K
Children of Color	
Lower staff reported *math development*	Infant/Toddler
Lower staff reported *language development*	Infant/Toddler & Preschool/Pre-K
Lower staff reported *social emotional development*	Infant/Toddler & Preschool/Pre-K
Lower staff reported *cognitive development*	Preschool/Pre-K
Lower staff reported *physical development*	Preschool/Pre-K
Lower parent reported *social emotional development*	Preschool/Pre-K
Higher parent reported *prosocial behavior*	Preschool/Pre-K
Child is Female	
Higher staff reported *language development*	Infant/Toddler
Higher parent reported *personal social skills*	Infant/Toddler
Fewer parent reported *peer problems*	Infant/Toddler
Lower staff reported *physical development*	Preschool/Pre-K
Higher parent reported *social emotional development*	Preschool/Pre-K
More parent reported *prosocial behavior*	Preschool/Pre-K
Fewer parent reported *conduct problems*	Preschool/Pre-K
Fewer parent reported *total difficulties*	Preschool/Pre-K
Online Version of the TS Gold (compared to Paper Version)	
Lower staff reported *language development*	Preschool/Pre-K
Lower staff reported *literacy development*	Preschool/Pre-K

Complete data tables can be found in Appendix C. Please note, although data are available in the tables for all 5 waves, there were many variables in Waves 3 through 5 that could not be accounted for (e.g., the effects of the COVID-19 pandemic on parents and on children; the fact that many children transitioned from one classroom to another and, due to the COVID-19 pandemic, the evaluation team was not able to collect data from the second classroom).

Parent Outcomes

Bivariate Correlations

Bivariate correlations are a simple measure of the relationship between two variables. These analyses include no other variables and are only used as preliminary analyses to inform subsequent analyses. Bivariate correlations of the parent outcomes and covariates at Wave 1 are presented in Table 65.

Table 65

Bivariate Correlations – Covariates and Parent Outcome Variables at Wave 1

	Parent Education	Number of Relocations	Years at Child Care Center	Parent Sex	Major Life Changes	Dual Military Status
Days Absent from Work in Last 6 Months Due to Child Care Arrangements	ns	.159*	ns	-.131*	ns	ns
Family-Work Conflict	ns	.159*	.134*	ns	ns	ns
Perceived Self-Efficacy	ns	156*	.144*	.128*	ns	ns
Perceived Helplessness	ns	.161*	ns	.143*	.150*	-.163*
Perceived Stress Total Scale	ns	.171**	ns	.147*	.140*	-.147*

Note. * p <.05, ** p<.01; ns = not significant. Parent Sex: male = 0; female = 1. Major Life Changes: no, did not experience a major life change = 0; yes, experienced a major life change = 1. Dual Military Status: not dual military = 0; yes, dual military = 1. Higher scores indicate more days absent, more conflict, less self-efficacy, more helplessness, and more total stress.

Work and Perceived Stress – Descriptive Statistics

Table 66 presents means for the parent outcomes: absenteeism from work, family-work conflict, and perceived stress. In initial analyses that compared the CDCs and the civilian centers on parent outcomes, with no covariates included, one difference emerged. Parents in the CDCs reported fewer days absent from work due to child care arrangements than did parents in civilian centers.

Table 66
Parent Work and Stress Means at Wave 1

	CDCs		Civilian Centers	
	Mean (SD)	n	Mean (SD)	n
Wave 1				
Days Absent from Work in Last 6 Months Due to Child Care Arrangements	2.30 (2.91)	189	4.34 (5.17)	37
Family-Work Conflict	2.39 (1.48)	190	2.50 (1.59)	37
Perceived Stress Scale - Efficacy	4.96 (2.79)	208	4.95 (2.77)	38
Perceived Stress Scale - Helplessness	8.94 (5.08)	208	9.79 (4.84)	38
Perceived Stress Scale - Total	13.90 (7.37)	208	14.74 (6.93)	38
Wave 3 (Cohort 2 during COVID)				
Days Absent from Work in Last 6 Months Due to Child Care Arrangements	-	-	-	-
Family-Work Conflict	2.72 (1.66)	154	2.69 (1.70)	31
Perceived Stress Scale - Efficacy	4.93 (3.00)	164	5.23 (2.64)	31
Perceived Stress Scale - Helplessness	8.77 (5.20)	164	8.48 (5.61)	31
Perceived Stress Scale - Total	13.71 (7.41)	164	13.71 (7.54)	31
Wave 4 (during COVID)				
Days Absent from Work in Last 6 Months Due to Child Care Arrangements	-	-	-	-
Family-Work Conflict	2.90 (1.77)	147	2.99 (1.96)	32
Perceived Stress Scale - Efficacy	5.35 (3.12)	161	5.00 (3.51)	32
Perceived Stress Scale - Helplessness	9.12 (5.34)	161	8.94 (6.41)	32
Perceived Stress Scale - Total	14.45 (7.66)	161	13.94 (9.40)	32
Wave 5 (during COVID)				
Days Absent from Work in Last 6 Months Due to Child Care Arrangements	-	-	-	-
Family-Work Conflict	2.82 (1.77)	137	2.94 (2.05)	28
Perceived Stress Scale - Efficacy	4.96 (3.17)	154	4.66 (3.73)	29
Perceived Stress Scale - Helplessness	8.92 (5.75)	154	8.62 (6.03)	29
Perceived Stress Scale - Total	13.88 (8.05)	154	13.28 (9.52)	29

Note. Blue font and red font indicate that there is a statistically significant difference ($p < .05$) between participants in the CDCs and the participants in the civilian centers on the indicated scale. Blue indicates a better outcome; red indicates a worse outcome. Higher scores indicate more days absent, more conflict, less self-efficacy, more helplessness, and more total stress. Differential attrition likely impacted perceived stress scores at Wave 4 and 5 as child well-being and parent stress are correlated.

Relation Between Early Childhood Education Factors and Parent Outcomes

The analyses of parent outcomes examined the relation between center type (i.e., CDC or civilian center); whether child care is available to the parent when the parent works late; whether child care is available when the child is mildly ill (i.e., pre-COVID-19); and the parent outcomes of absenteeism from work, family-work conflict, and perceived stress. The data summary is presented in Table 67. The following covariates were controlled for in the parent outcome analyses: parent education, number of relocations, years at current child care center, parent sex,

major life changes, and dual military status. The data summary for the covariates is presented in Table 68. The full data tables are available in Appendix C.

Table 67

Wave 1 Parent Outcomes

Child Availability – Working Late
There is less *absenteeism from work* when child care is available when parents need to work late

Table 68

Wave 1 Parent Covariates

More Relocations During Child's Life
More family-work conflict
More Years at Current Child Care Center
More family-work conflict
Less perceived self-efficacy
Female Respondent
More *total perceived stress*
More *perceived helplessness*
Dual Military Family
Less *perceived helplessness*

Limitations

Several threats to the internal validity[3] of this evaluation must be acknowledged.

History

In longitudinal designs, there is the possibility that events may happen that may impact the study or evaluation. There were several events that may have had an effect on the families involved in this evaluation.

- Hurricane Dorian impacted the southeast coast of the United States in September 2019. It affected the delivery of some parent questionnaires and the delivery of some Classroom Environment Questionnaires and TS Gold requests to CDCs. It also impacted the observer training.
- An airstrike was conducted in January of 2020 that resulted in the death of an Iranian general, which led to an 18-hour notice deployment for some Service members. This incident may have affected some of the families in this evaluation.
- The COVID-19 pandemic, which began disrupting American's day-to-day lives in March 2020, significantly disrupted multiple aspects of this evaluation; this disruption is described in detail on page 9.

Attrition

Although parent response rates for questionnaires that were distributed remained relatively high for the duration of the evaluation, there are some items that impacted the analyses.

- A significant number of eligible families were lost for known reasons. These reasons included the following:
    - 71 families did not complete the Wave 1 questionnaire;
        - This included 12 families for whom the evaluation team did not receive a permission form and, therefore, could not collect direct-care staff-reported child assessments.
        - Please note, the majority of parents who did not complete the Wave 1 questionnaire were from the same CDC.
    - 6 families from the civilian centers who did not complete Wave 1 and, therefore, had no data for the evaluation;
    - 28 families who disenrolled from the CDC; and
    - 5 families who withdrew from the evaluation.

[3] Internal validity is the extent to which a study establishes a trustworthy cause-and-effect relationship between a treatment and an outcome.

- In addition, as expected in longitudinal research and evaluation, there were parents who chose not to respond at some of the waves. At Wave 2, 35 parents chose not to respond; at Wave 3, 41 parents chose not to respond; at Wave 4, 50 parents chose not to respond; at Wave 5, 59 parents chose not to respond.
    - Note, COVID-19 likely influenced attrition at Waves 3, 4, and 5.

When attrition occurs, there is a possibility that there are differences between the families that leave the evaluation and families that remain in the evaluation. Indeed, there were differences between the families who attritted in this evaluation and those who did not. In particular, for children in the civilian centers, children in families that attritted had poorer well-being scores at Wave 1 than children in families who did not attrit. This was not the case for children in CDCs. Because of this difference in attrition, children in civilian centers appeared to do better on well-being over time than children in CDCs. However, through further analyses, the evaluation team was able to determine that this was not the case; it was due to the differential attrition.

Instrumentation

Several factors should be taken into consideration regarding the measures that were used in this analysis. In order reduce burden on CDC staff, this evaluation used the child assessment that the CDC staff were already using, the TS Gold. This measure was designed as a tool for early childhood education teachers to use to assess progress and guide instruction. As such, there are limitations for its use in evaluation. The analyses used in this evaluation were based on the widely held expectations categories developed by the measure developer. This constrains the analysis to a three-level outcome variable (i.e., 0, 1, and 2, to represent progressing towards expectations, meeting expectations, and exceeding expectations, respectively), whereas a continuous variable (e.g., a scale of 0-100) would be ideal for an outcome analysis. This same limitation also applies to the ASQ and ASQ:SE.

During the planning process, the evaluation team attempted to identify a parent-reported school readiness measure. No singular measure was identified. Therefore, a teacher-reported measure was adapted. As such, the results from the EDI should be interpreted with caution.

Sampling

Civilian Centers

Many of the civilian centers, which operated under multiple business names, were owned by the same corporation. The individual centers indicated that they could not agree to participate in the evaluation unless it was approved by the corporate office. Multiple attempts were made to get in touch with regional managers and the corporate headquarters, but email and phone communications sent by Penn State and OSD were never returned. Please note, this was not the

case for all corporate owners with a nationwide presence; some let their individual centers make the decision on involvement in the evaluation.

This resulted in a smaller than expected number of civilian centers from which recruitment could occur and a smaller number of families from civilian centers participating in the evaluation. This small sample size from civilian centers made comparisons between civilian centers and CDCs difficult.

CDCs

The Services chose which CDCs were to participate in this evaluation. They were provided with suggestions for choosing CDCs, which included larger programs, locations where large numbers of families were also using civilian care, and installations that were demographically representative of the military, as opposed to installations that largely accommodate new recruits, senior ranking individuals, or a particular set of job occupations. However, this was not a random sample of CDC programs, and, furthermore, randomly sampling CDCs within the services would likely not produce a demographically representative sample of children (i.e., due to some installations catering to particular job occupations, specialties, or paygrades). Thus, to what extent this evaluation can be generalized to all CDCs within the military is not clear. The evaluation team recommends caution with regards to over generalization.

Nested Data

The data collected for this evaluation are commonly referred to as nested data in the research community. That is, the data from respondents are not truly independent as children are nested within classrooms and families and teachers are nested within centers. This creates dependency in the data because individuals with shared experiences are more likely to provide similar responses on a survey than are individuals without shared experiences. Advanced statistical methods exist that can accommodate the dependency in the data, but the accuracy of these methods in estimating relationships among variables may be biased when sample sizes are small at the various levels of clustering (e.g., small number of children within a classroom). There are statistical corrections that can be applied to help counter the potential bias induced by small samples, but there is no absolute guarantee that these corrections will remove all bias.

Recommendations

Based on what was learned over the course of this evaluation, the evaluation team puts forth the following recommendations, which are consistent with the conclusions and recommendations in the National Academies of Sciences, Engineering, and Medicine's (2019) report on the military family readiness system. The recommendations encompass the following themes: (1) continuous quality improvement, (2) leveraging data and existing expertise, and (3) organizational climate.

Recommendation 1: In addition to the trainings currently offered to direct-care staff, implement professional development activities that are informed by this and future data collection.

- Use the specific indicators of the classroom quality assessment that were identified as areas for improvement (see pages 60 and 61).
- Conduct regular, unannounced classroom quality assessments to capture naturalistic data; employ these data for continuous quality improvement through professional development.
    - For this evaluation, the team had success hiring current and former military spouses to conduct these assessments.
- Implement additional monitoring by directors or training and curriculum specialists of the child assessment used by direct-care staff. Ensure correct categorization of children into class/grade and ensure assessments are completed in full and in compliance with training. When using any measure or tool, deviations from training commonly occur over time. When appropriate, conduct professional development to realign with training.
- Conduct yearly program- or Service-level analyses of collected TS Gold data and compare to TS Gold normed data. Use these data for continuous quality improvement through professional development.

Recommendation 2: Leverage the standards of state and national early childhood education organizations to inform decision-making.

- Align hiring practices with NAEYC education standards.

Recommendation 3: Monitor organizational climate and use evidence-informed practices and information from this evaluation and future data collection efforts to influence climate.

- To promote continuous quality monitoring, implement regular, externally conducted surveys of organizational climate and include questions related to director support and command support.
- Increase command support of the CDCs by increasing commanders' knowledge of the importance of high-quality child care for families, for the Service Member, and for mission readiness. Increase commanders' knowledge of the impact of command support on the child care centers. Ensure that command support is communicated effectively to program staff.

- Communicate bidirectional expectations of the staff-director relationship to both direct-care staff and the director. Use professional development activities to (1) ensure that directors are aware of what supports promote direct-care staff success and (2) ensure that direct-care staff have realistic expectations of director support.

Additional Considerations

Given the unanticipated challenges encountered during this evaluation (i.e., the COVID-19 pandemic) and the implementation of a new curriculum, additional evaluations would provide further, and more complete, data regarding military-supported early childhood education. An implementation evaluation would provide important information about how the new curriculum is operating on the ground, and an outcome evaluation would provide data that was disrupted by the COVID-19 pandemic.

- Implementation evaluation (ELM Curriculum)
    - Examine the frequency with which the curriculum is being used; whether it is being implemented with fidelity; if direct-care staff are experiencing any barriers to implementation; direct-care staff perception of the curriculum; and whether parents are satisfied with the curriculum.
- Outcome evaluation
    - Examine data longitudinally, over one year, as was originally intended.
    - Further explore the impact of direct-care staff education on classroom quality and child outcomes.
    - Investigate deviation from norms at Fall, Winter, and Spring TS Gold checkpoints.
    - The current evaluation is not generalizable to other curriculum implemented in the CDCs; to speak to the impact of the new ELM curriculum on child outcomes, an additional evaluation is required.
- Future evaluations should consider the following recommendations based on lessons learned from this evaluation:
    - In anticipation of unplanned disruptions to a longitudinal evaluation, carefully weigh the costs associated with collecting all measures at all timepoints (i.e., participant burden) against the benefits of having data for all measures at all timepoints should unexpected events occur.
    - To increase robustness of the metric, use a measure of child development that results in a continuous variable measuring development and is independent of the curriculum and assessments that are used by the CDCs.
    - Obtain approval to provide incentives to direct-care staff at the civilian centers in the evaluation, which participate in the DoD fee-assistance program, and collect the same data from the civilian centers as the data collected at the CDCs.
    - In an effort to potentially increase the number of participants from civilian centers, consider expanding the number of geographic locations from which civilian child care centers are recruited.
    - Proactively work with the DoD approvals offices to find a solution to allow the collection of child birthdates to ensure delivery of the correct versions of the measures.

References

Cohen, J. (1988*). Statistical power analysis for the behavioral sciences*. Routledge Academic.

Cohen, S., & Williamson, G. (1988). Perceived stress in a probability sample of the United States. In S. Spacapan & S. Oskamp (Eds.), *The social psychology of health* (pp. 31-67). Sage.

Davenport, K. E., Mathies, B. R., Karre, J. K., Kaye, M. P., & Perkins, D. F. (2016). *Child and Youth Development Evaluation Plan Project: Child Development Centers (CDC) and Family Child Care (FCC) evaluation plan.* Report submitted to the Department of Defense, Military Community and Family Policy. Clearinghouse for Military Family Readiness at Penn State.

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 1337-1245.

Goodman, R., Meltzer, H., & Bailey, V. (1998). The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *European Child and Adolescent Psychiatry*, *7*, 125-130.

Harms, T., Clifford, R. M., & Cryer, D. (2015). *Early Childhood Environment Rating Scale, third edition.* Teachers College Press.

Harms, T., Cryer, D., Clifford, R. M., & Yazejian, N. (2017). *Infant/Toddler Environmental Rating Scale, revised edition.* Teachers College Press.

Heroman, C., Burts, D., Berke, K., & Bickart, T. (2010). *Teaching Strategies GOLD objectives for development and learning*. Teaching Strategies.

Janus, M., & Offord, D. R. (2007). Development and psychometric properties of the Early Development Instrument (EDI): A measure of children's school readiness. *Canadian Journal of Behavioral Science, 39,* 1-22.

Kim, D. H., Lambert, R. G., & Burts, D. C. (2013). Evidence of the validity of Teaching Strategies GOLD assessment tool for English language learners and children with disabilities. *Early Education and Development, 24,* 574-595.

Lambert, R. G., (2020). *Technical manual for the Teaching Strategies GOLD Assessment* (2nd ed.). Teaching Strategies. https://teachingstrategies.com/wp-content/uploads/2020/09/2020-Tech-Report_GOLD_B-3_V4.pdf

Lambert, R. G., Kim, D. H., & Burts, D. C. (2015). The measurement properties of the Teaching Strategies GOLD® assessment system. *Early Childhood Research Quarterly, 33,* 49-63.

Lee, E. (2012). Review of the psychometric evidence of the perceived stress scale. *Asian Nursing Research, 6,* 121-127.

McNeish, D. M., & Harring, J. R. (2017). Clustered data with small sample sizes: Comparing the performance of model-based and design-based approaches. *Communications in Statistics – Simulation and Computation*, *46*, 855-869.

McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, *22*, 114-140.

Mellor, D. (2004). Furthering the use of the Strengths and Difficulties Questionnaire: Reliability with younger child respondents. *Psychological Assessment, 16,* 396-401.

National Academies of Sciences, Engineering, and Medicine. (2019). *Strengthening the military family readiness system for a changing American society.* The National Academies Press. https://doi.org/10.17226/25380

National Association for the Education of Young Children. (2019). *Professional standards and competencies for early childhood educators.* https://www.naeyc.org/sites/default/files/globally-shared/downloads/PDFs/resources/position-statements/standards_and_competencies_ps.pdf

Netemeyer, R. G., Boles, J. S., & McMurrian, R. (1996). Development and validation of work-family conflict and family-work conflict scales. *Journal of Applied Psychology, 81,* 400-410.

Roberti, J. W., Harrington, L. N., & Storch, E. A. (2006). Further psychometric support for the 10-item version of the perceived stress scale. *Journal of College Counseling*, *9*, 135-147.

Squires, J., & Bricker, D. (2009). *Ages & Stages Questionnaires®. A parent-completed child monitoring system* (3rd ed.). Paul Brookes Publishing Co.

Squires, J., Bricker, D., Heo, K., & Twombly, E. (2010). *Ages and Stages Questionnaires®. Social-Emotional. A parent completed, child-monitoring system for social emotional behaviors.* Paul H. Brookes Publishing Co.

Squires, J., Bricker, D., Twombly, E., Murphy, K., & Hoselton, R. (2015). *ASQ:SE-2 technical report.* Brookes Publishing. https://agesandstages.com/wp-content/uploads/2015/09/ASQSE-2-Technical-Appendix_web.pdf

Teaching Strategies. (2013*). Teaching Strategies GOLD® Assessment System concurrent validity. An independent study comparing external assessment measures to Teaching Strategies GOLD*. http://teachingstrategies.com/content/pageDocs/GOLD-Concurrent-Validity-2013.pdf

Appendices

Appendix A – Measure Reliability and Validity

ASQ-3

Reliability studies conducted on the ASQ-3 include test-retest reliability, interrater reliability, and internal consistency (Squires & Bricker, 2009). Parents completed questionnaires twice within a 2-week period to assess test-retest reliability. This resulted in 92% agreement and intraclass correlations from .75 to .82. Interrater reliability was assessed by comparing a parent's assessment of his or her child to an assessment completed by a trained examiner. The percent agreement was 93%, and intraclass correlations were .43 to .69. The lowest agreement was on the communication scale, and the highest agreement was on the personal-social scale. The scales demonstrated adequate internal consistency (α = .51 to .87). Validity was demonstrated by comparing assessments of at-risk to non-risk groups. Concurrent validity was examined by comparing children's scores on the ASQ-3 to another standardized test (i.e., the Battelle Developmental Inventory). Agreement between the two measures was .85. The ASQ-3 also demonstrates moderate sensitivity (.86) and specificity (.85).

ASQ:SE-2

Reliability studies conducted on the ASQ:SE include test-retest reliability and internal consistency (Squires et al., 2015). Parents completed questionnaires twice within a 2-week period to assess test-retest reliability. This resulted in 89% agreement and the intraclass correlation was .91. The scales demonstrate internal consistency (α = .84). Convergent validity of the ASQ:SE was examined by comparing children's scores on a number of standardized tests with social-emotional aspects (i.e., the Devereux Early Childhood Assessment for Infants and Toddlers, the Infant Toddler Social Emotional Assessment, and the Child Behavior Checklist). The ASQ:SE also demonstrates moderate sensitivity (.81) and specificity (.83).

EDI

Overall internal consistency for the EDI is satisfactory. For teacher reports of the physical health and well-being domain, the reliability is α = .84 ; the cognitive development domain reliability is α = .93; the communication skills domain is α = .95 (Janus & Offord, 2007). Parent reliability was not tested by Janus & Offord.

SDQ

The SDQ measure reliability is good with internal consistency that ranged from .57 to .85 (Goodman, 2001), and 2-week test-retest reliability mean coefficients ranged from .52 to .85 (Mellor, 2004). Although the coefficient for cross-informant reliability appears to be low, Goodman (2001) stated that it is higher than that of other measures. In addition to reliability, the SDQ was also found to be a valid instrument (Goodman, 2001).

FWC

The FWC scale has good internal consistency (α = .82-.90). There was also evidence of construct and discriminant validity (Netemeyer et al., 1996).

PSS-10

The PSS-10 has good internal consistency (α = .78), and there is evidence of construct, convergent, and divergent validity (Cohen & Williamson, 1988; Lee, 2012; Roberti et al., 2006).

TS GOLD

Reliability studies conducted on the TS GOLD include internal consistency and interrater reliability (Lambert et al., 2015), and the scales demonstrated good internal consistency (α = .93-.97). Interrater reliability was also examined. The correlations between the teacher and master rater on all scales were above .90 except the physical scale, which was .85. Concurrent validity was supported by comparing the TS GOLD to the Bracken School Readiness Assessment (Teaching Strategies, 2013). Validity was also established for children with disabilities and for children who speak English as a second language (Kim, Lambert, & Burts, 2013).

ITERS & ECERS

The scale has good overall internal consistency (α = .91; Harms et al., 2017). The internal consistency of the subscales ranges from .76 to .94. The ITERS also has good interrater reliability with intraclass correlations of .92 for the full scale and interclass correlations between .76 and .94 for the subscales.

The ECERS full scale has good internal consistency (α = .93; Harms et al., 2015). The internal consistency of the subscales ranges from .87 to .96. The ECERS also has good interrater reliability with intraclass correlations of .90 for the full scale and interclass correlations between .93 and .98 for the subscales.

Appendix B – Technical Appendix

Data Collection – Parent Report

- Parent questionnaires were sent within 1 week of the reported due date for the TS Gold direct-care staff assessment.
- Parent questionnaires expired 4 weeks after the survey was started or 8 weeks after the survey was sent, if they were never started.
- Cohort 1 Wave 2 and Cohort 2 Wave 1 surveys were distributed in the weeks leading up to Christmas 2019 and New Year's Day 2020. As such, questionnaire completion may have been affected by the timing. In an effort to mitigate the effects on the project, with the permission of OSD, a second reminder was sent out.
- With the permission of OSD, a third reminder was added after the onset of the COVID-19 pandemic.
- Six parents initially requested paper surveys. Four were completed, and two were not. Three of the four who completed the Wave 1 paper survey requested online surveys for subsequent data collections. The final participant who was completing paper surveys agreed to complete the survey online after the onset of the COVID-19 pandemic.
- Due to a technical issue, some registrants received the permission form at the beginning of Wave 1 (n=51) instead of at the time of registration (n=308).
- Two families in Cohort 1 did not complete the Wave 1 questionnaire before it expired, but they still wanted to participate; they were moved to Cohort 2.
- Cohort 1 Wave 5 and Cohort 2 Wave 4 parent data collection coincided with the opening of K-12 school for fall 2020. Many schools were engaging in remote learning, which would likely result in increased burden for many parents and may have affected the completion of or promptness of completion of questionnaires.
- Due to DoD regulations, the evaluation team was unable to collect child birthdates. Several of the measures had multiple versions based on the child's age. Therefore, the evaluation team created age ranges that the parent could select to ensure the correct version of the measure was received. In the questionnaire, the team included a link to an online calculator, so the parent could enter their child's birthdate and it would return the child's age in years, months, and days, which was consistent with the available age ranges to be selected. In addition, the email that was sent to the parent with the link to the questionnaire included the name of the child for whom the parent was completing the questionnaire. Despite the precautions taken, there were 35 families who, for at least one wave of data collection, selected an age range for their child that did not match the ages that were selected at other waves. It appeared that some parents selected the wrong age range, and other parents responded for the wrong child. For age-dependent measures, when the parent did not complete the correct version because they selected the incorrect age range, the evaluation team deleted responses. This resulted in at least some data being deleted for 32 families.

Data Collection – Observations

- Seven children transitioned from one classroom to another between the direct-care staff assessment and the first observation. The classroom that the child was in at the time of the observation is the observation score that was used in the analyses.
- Non-typical staff may have been in some classrooms during the observations. Some observers noted that the same staff member appeared at more than one classroom observation. Although it is possible that this individual may have been a floater, observers did not feel this was likely.
- Children may not have been physically present in the classroom when the observation occurred. They may have been absent on the day of the observation, or they may have transitioned to another classroom between the scheduling of the observation and the actual observation. However, the observation was of the classroom environment, not specifically of the child.

Data Collection – Direct-care Staff Report

- Classroom Environment Questionnaires and requests for TS Gold reports were sent within 1 month of the reported due date for the TS Gold direct-care staff assessments.
- For the seven children who transitioned from one classroom to another between the direct-care staff assessment and the first observation, the Classroom Environment Questionnaire was set to match the observation room.
- For several TS Gold reports, it appeared that children were listed in the wrong class/grade. The evaluation team identified 19 children at Wave 1, 21 children at Wave 2, and 2 children at Wave 3 who were not listed in the correct class/grade in TS Gold based on the director report of the children's classrooms, the parent report of the child's age, and the director report of the children who transitioned (or would have transitioned if not for COVID-19) to kindergarten the following fall. Because this affects the age band that the children are in, this, potentially, results in children being classified in the wrong category (i.e., approaching expectations, meeting expectations, exceeding expectations), so the evaluation team reclassified the class/grade for these children. This was only done when the team had confidence, based on the available data, that the child was listed in the incorrect class/grade. If there was any doubt, the child remained in the reported class/grade.
    - This appeared to occur for four reasons: (1) children were not moved to the Preschool 3 category after transitioning to the preschool classroom; (2) children were not moved into the Pre-K 4 category the fall before they would start kindergarten; (3) children were categorized as Pre-K 4 but should have been categorized as Preschool 3; and (4) children's birthdates appeared to be entered incorrectly.
    - It is possible that there were non-error reasons for the child being in the incorrect class/grade. (1) It could be that children were assessed while in Preschool 3 but transitioned to Pre-K 4 before the report was printed. (2) It could be that, although a child's age makes him or her eligible for kindergarten, he or she may not be developmentally ready for kindergarten and, thus, may stay in Pre-K a second year.

- The number of direct-care staff reported on the CEQ did not always match the number of direct-care staff reported in the classroom observations. This may have affected the direct-care staff education categorization. Forty-six classrooms had fewer direct-care staff listed on the CEQ than were listed on the observation form. Therefore, there may be incomplete data on direct-care staff education, which could result in lower categorizations than what is correct.

Disenrollment from the center and eligibility for the evaluation

- Prior to March 2020, if a child disenrolled from the child care center, he or she was considered ineligible to remain in the evaluation as he or she was no longer at a participating center. Therefore, the family no longer received questionnaires. However, beginning in March 2020, the evaluation team was unable to track which children were not in the centers due to COVID-19 closures, still in the centers as mission-essential families, enrolled in a different child care center as a stopgap, enrolled in a different center permanently, or left child care and remained at home. As such, if a parent asked if he or she should complete the questionnaire even though his or her child was not at the center at that time, he or she was advised to complete the survey.

Data Analysis

- Population Average Methods (PAMS) were used to account for the nested nature of the data (McNeish et al., 2017). These methods were selected over multilevel modeling (MLM) techniques because the primary interest was in estimating the fixed effects among predictor variables, covariates, and outcome variables, as opposed to estimating and explaining random effects. In general, PAMS and MLM will produce similar estimates of fixed effects, but PAMS can be more desirable because they require fewer assumptions. In addition, by not explicitly modeling random effects, convergence issues that may arise when certain data limitations are present can often be obviated.
- Two PAMS were specifically used for the evaluation: (1) generalized estimating equations (GEE) and (2) a restricted maximum likelihood estimation method. Both methods produce more appropriate standard errors than what would be obtained if the clustering were ignored (McNeish et al., 2017), which leads to more accurate assessments of statistical significance. For example, if standard errors were underestimated, then an effect may be deemed statistically significant when, in fact, it is not. The opposite problem is encountered if standard errors are overestimated. While both methods help to adjust the standard errors, they can result in different estimates of the fixed effects. This is because GEE does not use a maximum likelihood estimation technique to obtain fixed effects and because the GEE analysis process can modify the estimated effects in addition to the standard errors (McNeish et al., 2017).
- When applicable, the Kenward-Roger correction was employed to alleviate potential bias in the standard errors caused by small samples at the cluster level (e.g., classrooms, centers) or within the clusters (e.g., children in classrooms, teachers in centers). This correction can only be applied when using the restricted maximum likelihood estimation method.

Appendix C – Data Tables

Factors Related to Classroom Quality

Table 69

Associations between Direct-Care Staff Education and Classroom Quality (Continuous Variable) – ITERS (n=85) & ECERS (n=36)

	ITERS		ECERS	
	Unstandardized Coefficient	*p*	Unstandardized Coefficient	*p*
Space and Furnishings	.156	.098	-.012	.935
Personal Care Routines	.134	.031	-.077	.598
Language and Books/Literacy	.108	<.001	-.001	.995
Activities/Learning Activities	.116	.048	-.007	.964
Interaction	.122	.008	-.332	.214
Program Structure	.030	.650	.244	.310
Total Scale	.116	<.001	-.083	.577

Note. Blue font indicates statistical significance ($p < .05$). Statistical significance indicates that higher levels on the education variable are related to higher levels of quality on the subscale. Higher sores on the ITERS and ECERS indicate higher classroom quality.

Table 70

Association between Direct-Care Staff Education and Classroom Quality – ITERS & ECERS (n=121)

	At Least One Reported Bachelor's Degree	No Reported Bachelor's Degree	
	EMM (Std. Error)	EMM (Std. Error)	*p*
Space and Furnishings	5.35 (0.30)	4.80 (0.15)	.068
Personal Care Routines	4.48 (0.19)	4.32 (0.26)	.457
Language and Books/Literacy	4.79 (0.21)	4.40 (0.16)	.013
Activities/Learning Activities	3.79 (0.22)	3.35 (0.10)	.027
Interaction	5.25 (0.19)	5.06 (0.27)	.286
Program Structure	5.18 (0.21)	4.89 (0.32)	.196
Total Scale	4.66 (0.18)	4.33 (0.16)	.022

Note 1. EMM = Estimated Marginal Means (means adjusted for all other variables in the model). Blue and red indicate statistical significance ($p < .05$). This indicates that classrooms where at least one direct-care staff has a bachelor's degree (i.e., blue) have higher levels of quality on the subscale or total scale than classrooms where at least one direct-care staff does not have a bachelor's degree (i.e., red). Higher sores on the ITERS and ECERS indicate higher classroom quality.

Table 71

Association between Leadership Support and Classroom Quality – ITERS (n=82)

	ITERS			
	Director Support		Command Support	
	Unstandardized Coefficient	*p*	Unstandardized Coefficient	*p*
Space and Furnishings	-0.00 (0.07)	.987	0.13 (0.05)	.006
Personal Care Routines	-0.03 (0.14)	.825	0.14 (0.09)	.147
Language and Books	0.12 (0.09)	.202	0.21 (0.07)	.005
Activities	0.00 (0.11)	.981	0.08 (0.07)	.241
Interaction	0.18 (0.07)	.006	0.16 (0.08)	.039
Program Structure	0.29 (0.11)	.011	0.29 (0.09)	.002
Total Scale	0.08 (0.09)	.351	0.15 (0.06)	.007

Note. The three separate items and the two composite items were all highly correlated, so the effects were modeled in separate analyses but are shown in one table for ease of presentation. Blue indicates a statistically significant finding ($p < .05$) that is in the expected direction. This indicates that higher levels of support are related to higher levels of quality on the subscale or total scale. Higher sores on the ITERS and ECERS indicate higher classroom quality.

Table 72

Association between Leadership Support and Classroom Quality –ECERS (n=31)

	ECERS			
	Director Support		Command Support	
	Unstandardized Coefficient	*p*	Unstandardized Coefficient	*p*
Space and Furnishings	-0.10 (0.11)	.366	-0.14 (0.09)	.140
Personal Care Routines	-0.28 (0.20)	.152	-0.02 (0.17)	.890
Language and Literacy	-0.28 (0.17)	.093	-0.22 (0.18)	.223
Learning Activities	-0.11 (0.12)	.360	-0.16 (0.11)	.160
Interaction	-0.27 (0.19)	.150	-0.28 (0.20)	.163
Program Structure	-0.26 (0.23)	.252	-0.19 (0.21)	.386
Total Scale	-0.19 (0.10)	.069	-0.17 (0.11)	.143

Note. No statistically significant findings emerged. A lack of statistically significant findings does not necessarily mean that director and command support do not influence classroom quality for the Preschool/Pre-K rooms, it just means that a difference could not be detected. A small sample size for the Preschool/Pre-K classrooms may have affected significance testing. Higher sores on the ITERS and ECERS indicate higher classroom quality.

Correlations – Child Outcomes

Since data on the covariates were not collected at all waves, the wave for the covariate is either concurrent or the most recent past response. There are several instances in which results are significant at Waves 1 and 3 but not at Wave 2. It seems unlikely that this difference at Wave 2 is meaningful. It is more likely that this occurred due to a measurement issue. The evaluation team examined these data split by cohort and by those children who had data on Waves 1 and 2 and then for those who had data on Waves 1, 2, and 3. Cohort 1 seemed to be driving this lack of significance at Wave 2. This would have corresponded to the Fall checkpoint for this cohort. At the Fall checkpoint, children in Preschool and Pre-K are systematically moved into a new class/grade in the TS Gold system. The TS Gold normed data show that more children are in the Progressing Towards Expectations category at the Fall checkpoint than any other checkpoint. It is likely due to these measurement issues that the Wave 2 data diverge from Waves 1 and 3. Therefore, the evaluation team recommends focusing on Waves 1 and 3 when examining the TS Gold data.

Table 73.1

Bivariate Correlations – Covariates and Child Outcome Variables

		Child Sex	Children of Color (race)	Children of Color (ethnicity)	Parent Education	Parent Stress	Number of Relocations	Major Life Changes
TS Gold	Social-Emotional W1	ns	-.231**	ns	ns	ns	ns	ns
	Social-Emotional W2	ns	ns	ns	ns	ns	ns	ns
	Social-Emotional W3	ns	-.235*	ns	ns	ns	ns	ns
	Physical W1	ns	ns	ns	ns	ns	ns	ns
	Physical W2	ns	ns	ns	ns	ns	ns	ns
	Physical W3	ns	ns	ns	ns	ns	ns	ns
	Language W1	ns	-.257**	ns	ns	ns	ns	ns
	Language W2	ns	ns	ns	ns	ns	ns	ns
	Language W3	ns	-.207*	ns	ns	ns	ns	.229*
	Cognitive W1	ns	-.216**	-.177*	.159*	ns	ns	ns
	Cognitive W2	ns	ns	ns	ns	ns	ns	ns
	Cognitive W3	ns	-.215*	ns	.250*	ns	ns	ns
	Literacy W1	ns	ns	ns	.232**	ns	ns	ns
	Literacy W2	ns	ns	ns	ns	ns	ns	ns
	Literacy W3	ns	ns	ns	.335**	ns	ns	ns
	Math W1	ns	-.237**	ns	.198**	ns	ns	ns
	Math W2	ns	ns	ns	ns	ns	ns	ns
	Math W3	ns	ns	ns	ns	ns	ns	ns

Note. * p <.05, ** p < .01; ns = not significant. Child Sex: male = 0; female = 1. Children of Color (race): White = 0; Black or African American, American Indian or Alaska Native, or Asian = 1. Children of Color (ethnicity): No, not Spanish/Hispanic/Latino = 0; Yes, Mexican, Mexican-American, Chicano, Puerto Rican, Cuban, or other Spanish/Hispanic/Latino = 1. Major Life Change: No major life change in the last year = 0; Yes, major life change in the last year = 1. Higher scores on the TS Gold indicate more advanced levels of development.

Table 73.2

Bivariate Correlations – Covariates and Child Outcome Variables

		Child Sex	Children of Color (race)	Children of Color (ethnicity)	Parent Education	Parent Stress	Number of Relocations	Major Life Changes
ASQ-3	Communication W1	ns	.178**	ns	ns	ns	ns	ns
	Communication W2	ns	ns	ns	ns	ns	ns	ns
	Communication W3	-.177*	ns	ns	ns	ns	ns	ns
	Communication W4	ns	ns	ns	ns	ns	ns	ns
	Communication W5	ns	ns	ns	ns	-.176*	ns	ns
	Gross Motor W1	ns	ns	ns	ns	ns	ns	ns
	Gross Motor W2	ns	ns	ns	ns	ns	ns	ns
	Gross Motor W3	ns	ns	ns	ns	ns	ns	ns
	Gross Motor W4	ns	ns	ns	ns	ns	ns	ns
	Gross Motor W5	ns	ns	ns	ns	ns	ns	ns
	Fine Motor W1	ns	ns	ns	ns	ns	ns	ns
	Fine Motor W2	ns	ns	ns	ns	ns	ns	ns
	Fine Motor W3	ns	ns	ns	ns	.214**	ns	ns
	Fine Motor W4	ns	ns	ns	ns	ns	ns	ns
	Fine Motor W5	ns	ns	ns	ns	ns	ns	ns
	Problem Solving W1	ns	ns	ns	ns	ns	ns	ns
	Problem Solving W2	ns	ns	ns	ns	ns	ns	ns
	Problem Solving W3	ns	ns	ns	-.154*	ns	-.147*	ns
	Problem Solving W4	ns	ns	ns	ns	ns	ns	ns
	Problem Solving W5	ns	ns	ns	ns	ns	ns	ns
	Personal Social W1	-.161*	ns	ns	ns	ns	-.134*	ns
	Personal Social W2	ns	ns	ns	ns	ns	ns	ns
	Personal Social W3	ns	ns	ns	ns	ns	ns	ns
	Personal Social W4	ns	ns	ns	ns	ns	ns	ns
	Personal Social W5	-.170*	ns	ns	ns	ns	ns	ns

Note. * p <.05, ** p< .01; ns = not significant. Child Sex: male = 0; female = 1. Children of Color (race): White = 0; Black or African American, American Indian or Alaska Native, or Asian = 1. Children of Color (ethnicity): No, not Spanish/Hispanic/Latino = 0; Yes, Mexican, Mexican-American, Chicano, Puerto Rican, Cuban, or other Spanish/Hispanic/Latino = 1. Major Life Change: No major life change in the last year = 0; Yes, major life change in the last year = 1. Higher scores on the ASQ-3 indicate potential delayed development.

Table 73.3

Bivariate Correlations – Covariates and Child Outcome Variables

		Child Sex	Children of Color (race)	Children of Color (ethnicity)	Parent Education	Parent Stress	Number of Relocations	Major Life Changes
ASQ:SE	Social Emotional W1	-.135*	ns	ns	ns	.250**	ns	ns
	Social Emotional W2	-.163*	ns	ns	ns	.211**	ns	ns
	Social Emotional W3	-.178*	ns	ns	ns	ns	ns	ns
	Social Emotional W4	-.153*	.166*	ns	ns	.238**	ns	ns
	Social Emotional W5	ns	ns	ns	ns	.180*	ns	ns

Note. * p <.05, ** p < .01; ns = not significant. Child Sex: male = 0; female = 1. Children of Color (race): White = 0; Black or African American, American Indian or Alaska Native, or Asian = 1. Children of Color (ethnicity): No, not Spanish/Hispanic/Latino = 0; Yes, Mexican, Mexican-American, Chicano, Puerto Rican, Cuban, or other Spanish/Hispanic/Latino = 1. Major Life Change: No major life change in the last year = 0; Yes, major life change in the last year = 1. Higher scores on the ASQ:SE indicate potential delayed development.

Table 73.4

Bivariate Correlations – Covariates and Child Outcome Variables

		Child Sex	Children of Color (race)	Children of Color (ethnicity)	Parent Education	Parent Stress	Number of Relocations	Major Life Changes
SDQ 2-3	Emotional Symptoms W1	ns	ns	ns	ns	.238*	ns	ns
	Emotional Symptoms W3	ns	ns	ns	ns	ns	.278*	ns
	Emotional Symptoms W4	ns	ns	ns	ns	ns	ns	ns
	Emotional Symptoms W5	ns	ns	ns	ns	ns	.232*	ns
	Conduct Problems W1	ns	ns	.217*	ns	.363**	ns	ns
	Conduct Problems W3	ns	ns	ns	ns	.369**	ns	ns
	Conduct Problems W4	ns	ns	ns	ns	.466**	.268*	ns
	Conduct Problems W5	ns	ns	ns	ns	.436**	ns	ns
	Hyperactivity W1	-.237*	ns	ns	ns	ns	ns	ns
	Hyperactivity W3	-.257*	ns	ns	-.293*	ns	.318**	ns
	Hyperactivity W4	ns	ns	ns	ns	ns	ns	ns
	Hyperactivity W5	ns	ns	ns	ns	ns	ns	ns
	Peer Problems W1	-.242**	ns	.186*	ns	ns	.283**	ns
	Peer Problems W3	-.239*	ns	ns	ns	ns	.251*	ns
	Peer Problems W4	ns	ns	ns	ns	ns	ns	ns
	Peer Problems W5	ns	ns	ns	ns	ns	ns	ns
	Prosocial Behaviors W1	.290**	ns	ns	ns	ns	ns	ns
	Prosocial Behaviors W3	ns	ns	ns	ns	ns	ns	ns
	Prosocial Behaviors W4	ns	-.236*	ns	ns	ns	ns	ns
	Prosocial Behaviors W5	ns	ns	ns	ns	ns	ns	ns
	Internalizing W1	ns	ns	ns	ns	.223*	.226*	ns
	Internalizing W3	ns	ns	ns	ns	ns	.307**	ns
	Internalizing W4	ns	ns	ns	ns	ns	ns	ns
	Internalizing W5	ns	ns	-.272*	ns	ns	ns	ns
	Externalizing W1	-.198*	ns	ns	ns	.276**	ns	ns
	Externalizing W3	-.261*	ns	ns	-.255*	.318**	.272*	ns
	Externalizing W4	ns	ns	ns	ns	.385**	.296*	ns
	Externalizing W5	ns	ns	ns	ns	.347**	ns	ns
	Total Difficulties W1	-.217*	ns	ns	ns	.297**	.232*	ns
	Total Difficulties W3	-.258*	ns	ns	ns	.249*	.327**	ns
	Total Difficulties W4	ns	ns	ns	ns	.323**	.304**	ns
	Total Difficulties W5	-.222*	ns	ns	ns	.256*	ns	ns

Note. * p <.05, ** p < .01; ns = not significant. Child Sex: male = 0; female = 1. Children of Color (race): White = 0; Black or African American, American Indian or Alaska Native, or Asian = 1. Children of Color (ethnicity): No, not Spanish/Hispanic/Latino = 0; Yes, Mexican, Mexican-American, Chicano, Puerto Rican, Cuban, or other Spanish/Hispanic/Latino = 1. Major Life Change: No major life change in the last year = 0; Yes, major life change in the last year = 1. For all subscales except for Prosocial Behaviors, higher scores indicate more problematic behavior. For the Prosocial Behaviors subscale, higher scores indicate more prosocial behavior.

Table 73.5

Bivariate Correlations – Covariates and Child Outcome Variables

		Child Sex	Children of Color (race)	Children of Color (ethnicity)	Parent Education	Parent Stress	Number of Relocations	Major Life Changes
SDQ 4-5	Emotional Symptoms W1	ns	ns	ns	ns	.330*	ns	ns
	Emotional Symptoms W3	ns	ns	ns	ns	.377**	ns	ns
	Emotional Symptoms W4	ns	ns	ns	ns	.306**	ns	ns
	Emotional Symptoms W5	ns	ns	ns	ns	.473**	.251*	ns
	Conduct Problems W1	-.284*	ns	ns	ns	ns	ns	ns
	Conduct Problems W3	ns	ns	ns	ns	.330**	ns	ns
	Conduct Problems W4	ns	.238*	ns	ns	.248*	ns	ns
	Conduct Problems W5	ns	ns	ns	ns	.327**	ns	.293**
	Hyperactivity W1	ns	ns	ns	ns	ns	ns	ns
	Hyperactivity W3	ns	ns	ns	ns	ns	ns	ns
	Hyperactivity W4	ns	ns	ns	ns	.355**	ns	.307**
	Hyperactivity W5	ns	ns	ns	ns	.328**	.259*	.295**
	Peer Problems W1	ns	ns	ns	ns	ns	ns	ns
	Peer Problems W3	ns	ns	ns	ns	ns	ns	ns
	Peer Problems W4	ns	ns	ns	ns	.234*	ns	ns
	Peer Problems W5	ns	ns	ns	ns	.338**	ns	ns
	Prosocial Behaviors W1	ns	ns	ns	ns	ns	ns	.379**
	Prosocial Behaviors W3	ns	ns	ns	ns	-.266*	ns	ns
	Prosocial Behaviors W4	ns	ns	ns	ns	ns	ns	ns
	Prosocial Behaviors W5	ns	ns	.229*	ns	-.358**	ns	ns
	Internalizing W1	ns	ns	ns	ns	ns	ns	ns
	Internalizing W3	ns	ns	ns	ns	.356**	ns	ns
	Internalizing W4	ns	ns	ns	ns	.315**	ns	ns
	Internalizing W5	ns	ns	ns	ns	.463**	.229*	ns
	Externalizing W1	ns	ns	ns	ns	ns	ns	ns
	Externalizing W3	ns	ns	ns	ns	.296*	ns	ns
	Externalizing W4	ns	ns	ns	ns	.349**	ns	.296**
	Externalizing W5	ns	ns	ns	ns	.362**	ns	.325**
	Total Difficulties W1	ns	ns	ns	ns	ns	ns	ns
	Total Difficulties W3	ns	ns	ns	ns	.375**	ns	ns
	Total Difficulties W4	ns	ns	ns	ns	.396**	ns	.247*
	Total Difficulties W5	ns	ns	ns	ns	.480**	.258*	.241*

Note. * p <.05, ** p < .01; ns = not significant. Child Sex: male = 0; female = 1. Children of Color (race): White = 0; Black or African American, American Indian or Alaska Native, or Asian = 1. Children of Color (ethnicity): No, not Spanish/Hispanic/Latino = 0; Yes, Mexican, Mexican-American, Chicano, Puerto Rican, Cuban, or other Spanish/Hispanic/Latino = 1. Major Life Change: No major life change in the last year = 0; Yes, major life change in the last year = 1. For all subscales except for Prosocial Behaviors, higher scores indicate more problematic behavior. For the Prosocial Behaviors subscale, higher scores indicate more prosocial behavior.

Table 73.6

Bivariate Correlations – Covariates and Child Outcome Variables

		Child Sex	Children of Color (race)	Children of Color (ethnicity)	Parent Education	Parent Stress	Number of Relocations	Major Life Changes
SDQ 2-5	Emotional Symptoms W1	ns	ns	ns	ns	.260**	ns	ns
	Emotional Symptoms W3	ns	-.176*	ns	ns	.213*	.185*	ns
	Emotional Symptoms W4	ns	ns	ns	ns	.243**	ns	ns
	Emotional Symptoms W5	ns	ns	ns	ns	.344**	.280*	ns
	Conduct Problems W1	ns	ns	.155*	ns	.310**	ns	ns
	Conduct Problems W3	ns	ns	ns	ns	.358**	ns	.210*
	Conduct Problems W4	ns	ns	ns	ns	.363**	ns	ns
	Conduct Problems W5	ns	ns	ns	ns	.371**	ns	.212**
	Hyperactivity W1	ns	ns	ns	ns	ns	ns	ns
	Hyperactivity W3	-.203*	ns	ns	ns	.210*	.184*	ns
	Hyperactivity W4	-.170*	ns	ns	ns	.269**	.249**	ns
	Hyperactivity W5	-.177*	ns	ns	ns	.288**	.187*	.181*
	Peer Problems W1	ns	ns	ns	ns	ns	.196**	ns
	Peer Problems W3	ns	ns	ns	ns	ns	ns	ns
	Peer Problems W4	ns	ns	ns	ns	ns	ns	ns
	Peer Problems W5	ns	ns	ns	ns	ns	ns	ns
	Prosocial Behaviors W1	.253**	ns	ns	ns	-.182*	ns	ns
	Prosocial Behaviors W3	ns	ns	ns	ns	ns	ns	ns
	Prosocial Behaviors W4	ns	-.230**	ns	ns	ns	ns	ns
	Prosocial Behaviors W5	ns	ns	ns	ns	-.162*	ns	ns
	Internalizing W1	ns	ns	ns	ns	.230**	.183*	ns
	Internalizing W3	ns	ns	ns	ns	.215**	.176*	ns
	Internalizing W4	ns	ns	ns	ns	.221**	ns	ns
	Internalizing W5	ns	ns	ns	ns	.292**	.236**	ns
	Externalizing W1	-.162*	ns	ns	ns	.235**	ns	ns
	Externalizing W3	-.196*	ns	ns	ns	.310**	.166*	ns
	Externalizing W4	ns	ns	ns	ns	.360**	.219**	.207**
	Externalizing W5	-.176*	ns	ns	ns	.358**	.164*	.216**
	Total Difficulties W1	-.159*	ns	ns	ns	.272**	ns	ns
	Total Difficulties W3	ns	ns	ns	ns	.309**	.197*	ns
	Total Difficulties W4	ns	ns	ns	ns	.354**	.223**	ns
	Total Difficulties W5	-.171*	ns	ns	ns	.396**	.236**	.174*

Note. * p <.05, ** p < .01; ns = not significant. Child Sex: male = 0; female = 1. Children of Color (race): White = 0; Black or African American, American Indian or Alaska Native, or Asian = 1. Children of Color (ethnicity): No, not Spanish/Hispanic/Latino = 0; Yes, Mexican, Mexican-American, Chicano, Puerto Rican, Cuban, or other Spanish/Hispanic/Latino = 1. Major Life Change: No major life change in the last year = 0; Yes, major life change in the last year = 1. For all subscales except for Prosocial Behaviors, higher scores indicate more problematic behavior. For the Prosocial Behaviors subscale, higher scores indicate more prosocial behavior.

Table 73.7

Bivariate Correlations – Covariates and Child Outcome Variables

		Child Sex	Children of Color (race)	Children of Color (ethnicity)	Parent Education	Parent Stress	Number of Relocations	Major Life Changes
EDI	Language and Cognitive W1	ns	ns	ns	ns	ns	ns	ns
	Language and Cognitive W2	.310*	ns	ns	ns	-.368**	-.255*	ns
	Language and Cognitive W3	.251*	ns	ns	ns	-.247*	ns	ns
	Language and Cognitive W4	ns	ns	ns	ns	ns	ns	ns
	Language and Cognitive W5	ns	.234*	ns	ns	ns	ns	ns
	Physical W1	ns	ns	ns	.372**	-.326*	ns	ns
	Physical W2	ns	ns	ns	.251*	ns	ns	ns
	Physical W3	ns	ns	ns	ns	ns	ns	ns
	Physical W4	ns	.262*	ns	ns	ns	ns	ns
	Physical W5	ns	ns	ns	ns	-.275*	ns	ns
	Communication Skills W1	ns	ns	ns	ns	ns	.303*	ns
	Communication Skills W2	ns	ns	ns	ns	ns	ns	.278*
	Communication Skills W3	ns	ns	ns	.310**	ns	ns	ns
	Communication Skills W4	ns	ns	ns	ns	ns	ns	ns
	Communication Skills W5	.233*	ns	ns	ns	-.325**	ns	-.316**

Note. * p <.05, ** p < .01; ns = not significant. Child Sex: male = 0; female = 1. Children of Color (race): White = 0; Black or African American, American Indian or Alaska Native, or Asian = 1. Children of Color (ethnicity): No, not Spanish/Hispanic/Latino = 0; Yes, Mexican, Mexican-American, Chicano, Puerto Rican, Cuban, or other Spanish/Hispanic/Latino = 1. Major Life Change: No major life change in the last year = 0; Yes, major life change in the last year = 1. Higher scores indicate more advanced levels of development.

Factors Related to Child Outcomes

Analyses of child outcomes controlled for child sex, child race, child ethnicity, parent education, parent stress, number of relocations, and major life changes. For Waves 3 through 5, cohort was also included as a control variable. Furthermore, the TS Gold analyses for the language and literacy domains also included version (i.e., paper or online) as a covariate.

The original intent was to include each ITERS and ECERS subscale in the analyses to examine the association between each subscale and each outcome. However, the subscales were too correlated with each other to include simultaneously. Therefore, the first set of analyses included the total score. A second set of analyses was more targeted and included examining specific subscales that are likely to be related to specific outcomes, based on theory. For example, the relation between the Activities subscale and the Physical Development outcome was examined.

Please note, although data are available in the tables for all 5 waves, there were many external variables in Waves 3 through 5 that could not be accounted for (e.g., the effects of the COVID-19 pandemic on parents and children; the fact that many children transitioned from one classroom to another and, due to the COVID-19 pandemic, the evaluation team was not able to collect data

from the second classroom). Therefore, the reliability of these results is unclear, and decisions should not be made based on those results.

There are a few cases of counterintuitive results where alternative explanations are likely. For example, there was an instance where less director support appeared to be related to higher literacy development. However, since director support and command support are highly correlated, and since the estimates for director support and command support were the same or similar but in the opposite direction, it is likely that this occurred because of multicollinearity (i.e., the fact that they are highly correlated), and the negative result is unreliable. In addition, there were instances in which higher direct-care staff education was related to more emotional symptoms and more total difficulties. Based on prior research and theory, it is unlikely that higher direct-care staff education caused more emotional symptoms and total difficulties. It is more likely that center directors placed children with more emotional symptoms, and, by extension, more total difficulties, in classrooms where direct-care staff had more training or higher education. Finally, there were a few cases where, at later waves, higher classroom quality was related to worse outcomes. This is likely due to the demonstrated differential attrition or due to children transitioning to a different classroom.

Table 74

Factors Related to Child Development – TS Gold

	Social-Emotional	Physical	Language	Cognitive	Literacy	Math
			Wave 1			
Direct-Care Staff Education	ns	ns	ns	ns	ns	ns
Director Support	ns	ns	ns	ns	-0.18 (0.07)	ns
Command Support	ns	ns	ns	ns	0.18 (0.07)	0.15 (0.06)
Classroom Quality (ITERS)	ns	0.16 (0.06)	0.18 (0.06)	ns	ns	ns
Quality – Language & Books (ITERS)	-	-	ns	-	ns	-
Quality – Activities (ITERS)	-	0.19 (0.06)	-	ns	-	ns
Quality – Interaction (ITERS)	ns	-	-	-	-	-
Classroom Quality (ECERS)	ns	ns	ns	ns	ns	ns
Quality – Language & Literacy (ECERS)	-	-	ns	-	ns	-
Quality – Learning Activities (ECERS)	-	ns	-	ns	-	ns
Quality – Interaction (ECERS)	ns	-	-	-	-	-
			Wave 2			
Direct-Care Staff Education	ns	ns	ns	ns	ns	ns
Director Support	ns	ns	ns	ns	ns	ns
Command Support	ns	ns	ns	ns	ns	ns
Classroom Quality (ITERS)	ns	ns	ns	ns	ns	-0.18 (0.08)
Quality – Language & Books (ITERS)	-	-	ns	-	ns	-
Quality – Activities (ITERS)	-	ns	-	ns	-	ns
Quality – Interaction (ITERS)	ns	-	-	-	-	-
Classroom Quality (ECERS)	ns	ns	ns	ns	ns	ns
Quality – Language & Literacy (ECERS)	-	-	ns	-	ns	-
Quality – Learning Activities (ECERS)	-	ns	-	ns	-	ns
Quality – Interaction (ECERS)	ns	-	-	-	-	-
			Wave 3			
Direct-Care Staff Education	ns	ns	ns	ns	ns	ns
Director Support	ns	ns	ns	ns	ns	ns
Command Support	ns	ns	ns	ns	ns	ns
Classroom Quality (ITERS)	ns	ns	ns	ns	ns	ns
Quality – Language & Books (ITERS)	-	-	ns	-	ns	-
Quality – Activities (ITERS)	-	ns	-	ns	-	ns
Quality – Interaction (ITERS)	ns	-	-	-	-	-
Classroom Quality (ECERS)	ns	ns	ns	ns	ns	ns
Quality – Language & Literacy (ECERS)	-	-	ns	-	ns	-
Quality – Learning Activities (ECERS)	-	ns	-	ns	-	ns
Quality – Interaction (ECERS)	ns	-	-	-	-	-

Note. Estimates are unstandardized coefficients and standard errors for statistically significant results ($p < .05$). Blue font indicates better outcomes. Red font indicates worse outcomes. ns = not significant. Higher scores on the TS Gold indicate more advanced levels of development.

Table 75

Factors Related to Child Development - ASQ

	Communication	Gross Motor	Fine Motor	Problem Solving	Personal Social
Wave 1					
Center Type (Infant/Toddler) Center Type (Preschool/Pre-K)	ns	ns	ns	ns	-0.32 (0.15)[a] ns
Direct-Care Staff Education	ns	ns	ns	ns	ns
Director Support	ns	ns	ns	ns	ns
Command Support	ns	ns	ns	ns	ns
Classroom Quality (ITERS)	ns	ns	ns	-0.12 (0.06)	ns
Quality – Language & Books (ITERS)	ns	-	-	-	-
Quality – Activities (ITERS)	-	ns	ns	-0.13 (0.06)	-
Quality – Interaction (ITERS)	-	-	-	-	ns
Classroom Quality (ECERS)	ns	ns	ns	ns	ns
Quality – Language & Literacy (ECERS)	ns	-	-	-	-
Quality – Learning Activities (ECERS)	-	ns	ns	ns	-
Quality – Interaction (ECERS)	-	-	-	-	ns
Wave 2					
Center Type	ns	ns	ns	ns	ns
Direct-Care Staff Education	ns	ns	ns	ns	ns
Director Support	ns	ns	ns	ns	ns
Command Support	ns	ns	ns	ns	ns
Classroom Quality (ITERS)	ns	ns	ns	ns	ns
Quality – Language & Books (ITERS)	ns	-	-	-	-
Quality – Activities (ITERS)	-	ns	ns	ns	-
Quality – Interaction (ITERS)	-	-	-	-	ns
Classroom Quality (ECERS)	ns	ns	ns	ns	ns
Quality – Language & Literacy (ECERS)	ns	-	-	-	-
Quality – Learning Activities (ECERS)	-	ns	ns	ns	-
Quality – Interaction (ECERS)	-	-	-	-	ns
Wave 3 (Cohort 2 during COVID)					
Center Type (Infant/Toddler) Center Type (Preschool/Pre-K)	ns	ns	-0.26 (0.11) ns	ns	ns
Direct-Care Staff Education	ns	ns	ns	ns	ns
Director Support	ns	ns	ns	ns	ns
Command Support	ns	ns	ns	ns	ns
Classroom Quality (ITERS)	ns	ns	ns	ns	ns
Quality – Language & Books (ITERS)	ns	-	-	-	-
Quality – Activities (ITERS)	-	ns	ns	ns	-
Quality – Interaction (ITERS)	-	-	-	-	ns
Classroom Quality (ECERS)	ns	ns	ns	ns	ns
Quality – Language & Literacy (ECERS)	ns	-	-	-	-
Quality – Learning Activities (ECERS)	-	ns	ns	ns	-
Quality – Interaction (ECERS)	-	-	-	-	ns

continued

Table 75 (continued)

Factors Related to Child Development – ASQ

	Communication	Gross Motor	Fine Motor	Problem Solving	Personal Social
Wave 4 (during COVID)					
Center Type	ns	ns	ns	ns	ns
Direct-Care Staff Education	ns	ns	ns	ns	ns
Director Support	ns	ns	ns	ns	ns
Command Support	ns	ns	ns	ns	ns
Classroom Quality (ITERS)	-0.09 (0.04)	ns	ns	ns	-0.10 (0.05)
Quality - Language & Books (ITERS)	-0.09 (0.03)	-	-	-	-
Quality – Activities (ITERS)	-	ns	ns	ns	-
Quality – Interaction (ITERS)	-	-	-	-	ns
Classroom Quality (ECERS)	No converge	ns	ns	No converge	No converge
Quality - Language & Literacy (ECERS)	No converge	-	-	-	-
Quality – Learning Activities (ECERS)	-	ns	ns	No converge	-
Quality – Interaction (ECERS)	-	-	-	-	No converge
Wave 5 (during COVID)					
Center Type	ns	ns	ns	ns	ns
Direct-Care Staff Education	ns	ns	ns	ns	ns
Director Support	ns	ns	ns	ns	ns
Command Support	ns	ns	ns	ns	ns
Classroom Quality (ITERS)	-0.15 (0.05)	ns	ns	ns	ns
Quality - Language & Books (ITERS)	-0.11 (0.04)	-	-	-	-
Quality – Activities (ITERS)	-	ns	ns	ns	-
Quality – Interaction (ITERS)	-	-	-	-	ns
Classroom Quality (ECERS)	No converge	ns	ns	No converge	ns
Quality - Language & Literacy (ECERS)	No converge	-	-	-	-
Quality – Learning Activities (ECERS)	-	ns	ns	No converge	-
Quality – Interaction (ECERS)	-	-	-	-	ns

Note. No converge indicates that there was not sufficient variability in the outcome variable for the statistical model to run the model. Estimates are unstandardized coefficients and standard errors for statistically significant results ($p < .05$). Blue font indicates better outcomes. ns = not significant. Higher scores on the ASQ-3 indicate potential delayed development.

[a] In the analysis of the targeted Interaction ITERS subscale and the Personal Social outcome, center type emerged as significant. In the model with the total scale, that coefficient was almost equal (.30) and was approaching significant ($p = .06$). Due to this convergence, it is included as significant in the data summary.

Table 76

Factors Related to Child Development – ASQ:SE

	Social Emotional Development
Wave 1	
Center Type (Infant/Toddler)	ns
Center Type (Preschool/Pre-K)	-0.38 (0.17)
Direct-Care Staff Education	ns
Director Support	ns
Command Support	ns
Classroom Quality (ITERS)	ns
Quality – Interaction (ITERS)	ns
Classroom Quality (ECERS)	ns
Quality – Interaction (ECERS)	ns
Wave 2	
Center Type	ns
Direct-Care Staff Education	ns
Director Support	ns
Command Support	ns
Classroom Quality (ITERS)	ns
Quality – Interaction (ITERS)	0.10 (0.03)
Classroom Quality (ECERS)	ns
Quality – Interaction (ECERS)	ns
Wave 3 (Cohort 2 during COVID)	
Center Type	ns
Direct-Care Staff Education	ns
Director Support	ns
Command Support	ns
Classroom Quality (ITERS)	ns
Quality – Interaction (ITERS)	ns
Classroom Quality (ECERS)	ns
Quality – Interaction (ECERS)	ns
Wave 4 (during COVID)	
Center Type	ns
Direct-Care Staff Education	ns
Director Support	ns
Command Support	ns
Classroom Quality (ITERS)	ns
Quality – Interaction (ITERS)	ns
Classroom Quality (ECERS)	ns
Quality – Interaction (ECERS)	ns

continued

Table 76 (continued)

Factors Related to Child Development – ASQ:SE

	Social Emotional Development
Wave 5 (during COVID)	
Center Type	ns
Direct-Care Staff Education	ns
Director Support	ns
Command Support	ns
Classroom Quality (ITERS)	ns
Quality – Interaction (ITERS)	ns
Classroom Quality (ECERS)	ns
Quality – Interaction (ECERS)	ns

Note. Estimates are unstandardized coefficients and standard errors for statistically significant results ($p < .05$). Blue font indicates better outcomes. Red font indicates worse outcomes. ns = not significant. Higher scores on the ASQ-3 indicate potential delayed development.

Table 77

Factors Related to Child Well-Being – SDQ Bands

	Emotional Symptoms	Conduct Problems	Hyperactivity	Peer Problems	Prosocial Behaviors	Total Difficulties
Wave 1						
Center Type (Infant/Toddler) Center Type (Preschool/Pre-K)	ns	ns -0.57 (0.19)	ns	ns	ns	ns -0.40 (0.20)
Direct-Care Staff Education	0.14 (0.05)	ns	ns	ns	ns	0.12 (0.04)
Director Support	ns	-0.20 (0.08)	ns	ns	ns	ns
Command Support	ns	ns	ns	ns	ns	ns
Classroom Quality (ITERS)	ns	ns	ns	ns	ns	ns
Quality – Interaction (ITERS)	ns	ns	ns	ns	ns	ns
Classroom Quality (ECERS)	ns	ns	ns	ns	ns	ns
Quality – Interaction (ECERS)	ns	ns	ns	ns	ns	ns
Wave 3 (Cohort 2 during COVID)						
Center Type	ns	ns	ns	ns	ns	ns
Direct-Care Staff Education	ns	ns	ns	ns	ns	ns
Director Support	ns	-0.30 (0.07)	ns	ns	ns	ns
Command Support	ns	0.19 (0.07)	ns	ns	ns	ns
Classroom Quality (ITERS)	ns	ns	ns	ns	ns	ns
Quality – Interaction (ITERS)	ns	ns	ns	ns	ns	ns
Classroom Quality (ECERS)	ns	ns	ns	ns	ns	ns
Quality – Interaction (ECERS)	ns	ns	ns	ns	ns	ns
Wave 4 (during COVID)						
Center Type	ns	ns	ns	ns	ns	ns
Direct-Care Staff Education	ns	ns	ns	ns	ns	ns
Director Support	ns	ns	ns	ns	ns	ns
Command Support	ns	ns	ns	ns	ns	ns
Classroom Quality (ITERS)	ns	ns	ns	0.21 (0.08)	ns	ns
Quality – Interaction (ITERS)	ns	ns	ns	0.13 (0.05)	ns	ns
Classroom Quality (ECERS)	ns	ns	ns	ns	-0.31 (0.11)	ns
Quality – Interaction (ECERS)	ns	ns	ns	ns	-0.16 (0.07)	ns
Wave 5 (during COVID)						
Center Type	ns	ns	ns	ns	ns	ns
Direct-Care Staff Education	ns	ns	ns	ns	ns	ns
Director Support	ns	ns	ns	ns	ns	ns
Command Support	ns	ns	ns	ns	ns	ns
Classroom Quality (ITERS)	ns	ns	ns	ns	0.19 (0.07)	ns
Quality – Interaction (ITERS)	ns	ns	ns	ns	0.10 (0.05)	ns
Classroom Quality (ECERS)	ns	ns	ns	ns	ns	ns
Quality – Interaction (ECERS)	ns	ns	ns	ns	ns	ns

Note. Estimates are unstandardized coefficients and standard errors for statistically significant results ($p < .05$). Blue font indicates better outcomes. Red font indicates worse outcomes. ns = not significant. For all subscales, lower scores indicate scores that are closer to average; elevated scores indicate more problematic behavior.

Table 78

Factors Related to Child Well-Being – SDQ Means

	Emotional Symptoms	Conduct Problems	Hyperactivity	Peer Problems	Prosocial Behaviors	Total Difficulties
Wave 1						
Center Type (Infant/Toddler) Center Type (Preschool/Pre-K)	ns	ns -1.05 (0.50)	ns	ns	ns	ns
Direct-Care Staff Education	.22 (0.10)	ns	ns	ns	ns	ns
Director Support	ns	ns	ns	ns	ns	ns
Command Support	ns	ns	ns	ns	ns	ns
Classroom Quality (ITERS)	ns	ns	ns	ns	ns	ns
Quality – Interaction (ITERS)	ns	ns	ns	ns	ns	ns
Classroom Quality (ECERS)	ns	ns	ns	ns	ns	ns
Quality – Interaction (ECERS)	ns	ns	ns	ns	ns	ns
Wave 3 (Cohort 2 during COVID)						
Center Type	ns	ns	ns	ns	ns	ns
Direct-Care Staff Education	ns	ns	ns	ns	ns	ns
Director Support	ns	-0.63 (0.20)	ns	ns	ns	ns
Command Support	ns	0.50 (0.19)	ns	ns	ns	ns
Classroom Quality (ITERS)	ns	ns	ns	ns	ns	ns
Quality – Interaction (ITERS)	ns	ns	ns	ns	ns	ns
Classroom Quality (ECERS)	ns	ns	ns	ns	ns	ns
Quality – Interaction (ECERS)	ns	ns	ns	ns	ns	ns
Wave 4 (during COVID)						
Center Type (Infant/Toddler) Center Type (Preschool/Pre-K)	ns 1.68 (0.66)	ns	1.43 (0.52) 1.84 (0.85)	ns	ns	ns 5.06 (1.89)
Direct-Care Staff Education	ns	ns	ns	0.32 (0.13)	ns	ns
Director Support	ns	-0.48 (0.22)	ns	ns	ns	ns
Command Support	ns	ns	ns	ns	ns	ns
Classroom Quality (ITERS)	ns	ns	ns	0.48 (0.16)	ns	ns
Quality – Interaction (ITERS)	ns	ns	ns	0.30 (0.10)	ns	ns
Classroom Quality (ECERS)	ns	ns	ns	ns	0.53 (0.23)	ns
Quality – Interaction (ECERS)	ns	ns	ns	ns	ns	ns
Wave 5 (during COVID)						
Center Type (Infant/Toddler) Center Type (Preschool/Pre-K)	ns 1.79 (0.61)	ns	ns	ns	ns	ns 3.99 (1.83)
Direct-Care Staff Education	ns	ns	ns	ns	ns	ns
Director Support	ns	ns	ns	ns	ns	ns
Command Support	ns	ns	ns	ns	ns	ns
Classroom Quality (ITERS)	ns	ns	ns	0.34 (0.16)	ns	ns
Quality – Interaction (ITERS)	ns	ns	ns	ns	ns	ns
Classroom Quality (ECERS)	ns	-0.39 (0.15)	ns	ns	0.56 (0.18)	ns
Quality – Interaction (ECERS)	ns	-0.20 (0.09)	ns	ns	0.31 (0.11)	ns

Note. Estimates are unstandardized coefficients and standard errors for statistically significant results ($p < .05$). Blue font indicates better outcomes. Red font indicates worse outcomes. Brown font indicates a statistically significant outcome that can be explained by differential attrition. ns = not significant. For all subscales except for Prosocial Behaviors, higher scores indicate more problematic behavior. For the Prosocial Behaviors subscale, higher scores indicate more prosocial behavior.

Table 79

Factors Related to Child Development - EDI

	Language and Cognitive Development	Physical Well-being	Communication Skills
Wave 1			
Center Type	ns	ns	ns
Direct-Care Staff Education	ns	ns	ns
Director Support	ns	ns	ns
Command Support	ns	ns	ns
Classroom Quality (ECERS)	ns	ns	ns
Quality - Language & Literacy (ECERS)	ns	-	ns
Quality – Learning Activities (ECERS)	ns	ns	ns
Wave 2			
Center Type	ns	ns	ns
Direct-Care Staff Education	ns	ns	ns
Director Support	ns	ns	ns
Command Support	ns	ns	ns
Classroom Quality (ECERS)	ns	ns	ns
Quality - Language & Literacy (ECERS)	ns	-	ns
Quality – Learning Activities (ECERS)	ns	ns	ns
Wave 3 (Cohort 2 during COVID)			
Center Type	ns	ns	ns
Direct-Care Staff Education	ns	ns	ns
Director Support	ns	ns	ns
Command Support	ns	ns	ns
Classroom Quality (ECERS)	ns	ns	ns
Quality - Language & Literacy (ECERS)	ns	-	ns
Quality – Learning Activities (ECERS)	ns	ns	ns
Wave 4 (during COVID)			
Center Type	ns	ns	ns
Direct-Care Staff Education	ns	ns	ns
Director Support	ns	ns	ns
Command Support	ns	ns	ns
Classroom Quality (ECERS)	ns	ns	ns
Quality – Language & Literacy (ECERS)	ns	-	ns
Quality – Learning Activities (ECERS)	ns	ns	ns
Wave 5 (during COVID)			
Center Type	ns	ns	ns
Direct-Care Staff Education	ns	ns	ns
Director Support	ns	ns	ns
Command Support	ns	ns	ns
Classroom Quality (ECERS)	ns	ns	ns
Quality - Language & Literacy (ECERS)	ns	-	ns
Quality – Learning Activities (ECERS)	ns	ns	ns

Note. ns = not significant. Higher scores indicate more advanced levels of development.

Correlations – Parent Outcomes

Table 80

Bivariate Correlations – Covariates and Child Outcome Variables

	Parent Education	Number of Relocations	Years at Child Care Center	Parent Sex	Major Life Changes	Dual Military Status
Days Absent from Work in Last 6 Months Due to Child Care Arrangements W1	ns	.159*	ns	-.131*	ns	ns
Family-Work Conflict W1	ns	.159*	.134*	ns	ns	ns
Family-Work Conflict W3	.247**	ns	ns	ns	.170*	ns
Family-Work Conflict W4	.179*	.171*	ns	ns	.258**	ns
Family-Work Conflict W5	.203**	.283**	ns	ns	.312**	ns
Perceived Self-Efficacy W1	ns	156*	.144*	.128*	ns	ns
Perceived Self-Efficacy W3	ns	ns	ns	ns	ns	ns
Perceived Self-Efficacy W4	ns	.171*	ns	ns	ns	ns
Perceived Self-Efficacy W5	ns	.173*	ns	ns	ns	ns
Perceived Helplessness W1	ns	.161*	ns	.143*	.150*	-.163*
Perceived Helplessness W3	ns	ns	.150*	.217**	.251**	-.163*
Perceived Helplessness W4	ns	.194**	ns	ns	.192**	ns
Perceived Helplessness W5	ns	.229**	ns	ns	.233**	ns
Perceived Stress Total Scale W1	ns	.171**	ns	.147*	.140*	-.147*
Perceived Stress Total Scale W3	ns	ns	ns	.193**	.229**	-.166*
Perceived Stress Total Scale W4	ns	.202**	ns	ns	.177*	ns
Perceived Stress Total Scale W5	ns	.228**	ns	ns	.189*	ns

Note. * p <.05, ** p < .01; ns = not significant. Parent Sex: male = 0; female = 1. Major Life Change: No major life change in the last year = 0; Yes, major life change in the last year = 1. Dual Military Status: not dual military = 0; yes, dual military = 1. Higher scores indicate more days absent, more conflict, less self-efficacy, more helplessness, and more total stress.

Factors Related to Parent Outcomes

Table 81

Factors Related to Parent Outcomes

	Absenteeism From Work	Family-Work Conflict	Perceived Stress	Stress – Helplessness	Stress – Self-Efficacy
Wave 1					
Center Type	ns	ns	ns	ns	ns
Child care not available – work late	1.20 (0.53)	ns	ns	ns	ns
Child care not available – child is ill	ns	ns	ns	ns	ns
Wave 3 (Cohort 2 during COVID)					
Center Type	-	ns	ns	ns	ns
Child care not available – work late	-	0.73 (0.26)	3.63 (1.13)	2.83 (0.77)	ns
Child care not available – child is ill	-	ns	ns	ns	ns
Wave 4 (during COVID)					
Center Type	-	ns	ns	ns	ns
Child care not available – work late	-	ns	ns	ns	ns
Child care not available – child is ill	-	ns	ns	ns	1.25 (0.61)
Wave 5 (during COVID)					
Center Type	-	ns	ns	ns	ns
Child care not available – work late	-	0.97 (0.30)	3.93 (1.50)	3.32 (1.01)	ns
Child care not available – child is ill	-	ns	ns	ns	ns

Note. Estimates are unstandardized coefficients and standard errors for statistically significant results ($p < .05$). Blue font indicates better outcomes. ns = not significant. Higher scores indicate more days absent, more conflict, less self-efficacy, more helplessness, and more total stress.

Appendix D – Lessons Learned

- DoD approvals take time. Approvals for this project took over a year; conducting the evaluation as program evaluation as opposed to research did not speed up the approvals process.
- Civilian centers, as a whole, were not eager to participate; some needed approval at the corporate level.
- Civilian centers reported higher numbers of military families than was reported on the list from Child Care Aware of America.
- Per DoDI 3216.02, incentives can be provided to Service members, spouses, and civilian DoD employees.
- Incentivizing civilians who were not DoD employees would require Office of Management and Budget (OMB) approval (~2 years).
    - Stakeholders originally requested to compare the military children in the CDCs to civilian children in civilian care. However, incentivizing civilians who are not DoD employees would require OMB approval, which would take 2 years. Therefore, the comparison group was changed to military children in civilian care.
    - However, this barrier still impacted the evaluation because the direct-care staff at the civilian centers could not be incentivized, and, therefore, no direct-care staff-reported data (e.g., direct-care staff-reported child assessment, direct-care staff-education level) was collected from the civilian centers.
    - Similarly, approval was not granted to incentivize (e.g., book voucher, gift card to educational supply store) civilian centers with DoD funds.
- The evaluation team was not allowed to collect actual birthdates, therefore parents had to select from 27 choices for the age of their 0- to 5-year-old in order to receive the correct versions of the questionnaires.
    - Although a link to an online birthdate calculator was provided that allowed parents to convert their child's birthdate into their child's age in years, months, and days, selecting the child's age from a dropdown menu of 27 choices appeared to be a challenge for some parents. The evaluation team found 35 instances in which parents clearly selected the incorrect age range (e.g., the child became younger at a subsequent data-collection timepoint). Finding an acceptable way to collect children's actual birthdates will be important in future endeavors that collect data using age-dependent measures.
- Asking parents their child's room name is not a reliable way to get that information; 25% of parents listed incorrect or incomplete classroom names.