



UvA-DARE (Digital Academic Repository)

Child care quality in The Netherlands over the years: a closer look

Helmerhorst, K.O.W.; Riksen-Walraven, J.M.A.; Gevers Deynoot-Schaub, M.J.J.M.; Tavecchio, L.W.C.; Fukkink, R.G.

DOI

[10.1080/10409289.2014.948784](https://doi.org/10.1080/10409289.2014.948784)

Publication date

2015

Document Version

Final published version

Published in

Early Education and Development

License

Article 25fa Dutch Copyright Act

[Link to publication](#)

Citation for published version (APA):

Helmerhorst, K. O. W., Riksen-Walraven, J. M. A., Gevers Deynoot-Schaub, M. J. J. M., Tavecchio, L. W. C., & Fukkink, R. G. (2015). Child care quality in The Netherlands over the years: a closer look. *Early Education and Development*, 26(1), 89-105.
<https://doi.org/10.1080/10409289.2014.948784>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

Child Care Quality in The Netherlands Over the Years: A Closer Look

Katrien O. W. Helmerhorst

*Research Institute of Child Development and Education,
University of Amsterdam*

J. Marianne A. Riksen-Walraven

Department of Developmental Psychology, Radboud University Nijmegen

Mirjam J. J. M. Gevers Deynoot-Schaub

The Kohnstamm Institute, University of Amsterdam

Louis W. C. Tavecchio and Ruben G. Fukkink

*Research Institute of Child Development and Education,
University of Amsterdam*

Research Findings: We assessed the quality of child care in a nationally representative sample of 200 Dutch child care centers using the Infant/Toddler Environment Rating Scale–Revised and/or Early Childhood Environment Rating Scale–Revised and compared it with a previous assessment in 2005. The Caregiver Interaction Profile (CIP) scales were used to rate the quality of caregiver–child interactions. Results showed a significant and substantial decline in quality compared to 2005, with 49% of the groups now scoring below the minimal level. The CIP scales showed relatively high scores for the basic caregiver interactive skills of sensitive responsiveness, respect for autonomy, and structuring and limit setting but much lower scores for the more educational skills of verbal communication, developmental stimulation, and fostering positive peer interactions. Caregiver sensitive responsiveness was significantly lower in infant groups (0–2 years) than in preschool groups (2–4 years); caregiver respect for autonomy, verbal communication, developmental stimulation, and fostering positive peer interactions were significantly lower in infant groups than in preschool groups and mixed-age groups (0–4 years). *Practice or Policy:* Quality of child care is not stable across the years, and regular quality assessments are therefore needed to monitor child care quality. The low scores on the more educational versus the more basic caregiver interactive skills indicate that these skills deserve more attention in caregiver education and training. Training programs should be attuned to the individual interaction skill profile of caregivers.

Correspondence regarding this article should be addressed to Katrien O. W. Helmerhorst, Research Institute of Child Development and Education, University of Amsterdam, P.O. Box 15780, 1001 NG Amsterdam, The Netherlands. E-mail: k.o.w.helmerhorst@uva.nl

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/heed.

Quality of early child care can be broadly defined as the extent to which the child care environment offers experiences to children that enhance their well-being and development (see, e.g., Layzer & Goodson, 2006). There is now ample empirical evidence that high-quality child care may have considerable and enduring effects on children's cognitive and socioemotional development (for an overview, see Belsky et al., 2007; Vandell et al., 2010; Vandell & Wolfe, 2000). It is therefore important and in the interest of children to closely monitor the quality of care and improve it when needed.

In The Netherlands, repeated quality assessments in nationally representative samples of child care centers have shown a steady decrease in average quality from 1995 to 2001 and 2005 (Gevers Deynoot-Schaub & Riksen-Walraven, 2005; Van IJzendoorn, Tavecchio, Stams, Verhoeven, & Reiling, 1998; Vermeer et al., 2008). In these Dutch studies, child care quality was assessed using the Infant/Toddler Environment Rating Scale–Revised (ITERS-R; Harms, Cryer, & Clifford, 2003) and its equivalent for preschool classrooms, the Early Childhood Environment Rating Scale–Revised (ECERS-R; Harms, Clifford, & Cryer, 1998). These scales are used to rate the global quality of the child care environment in different domains, such as space and furnishings, personal care routines, language, activities, interactions, and program structure. In 1995, this instrument was used for the first Dutch child care quality assessment (Van IJzendoorn et al., 1998), because it was used in many countries and therefore allowed for an international comparison (e.g., Tietze, Cryer, Bairrão, Palacios, & Wetzel, 1996). In later national quality assessments, we retained the ITERS-R and ECERS-R as measures to be able to compare global quality ratings over time. The present study reports the results of the fourth Dutch quality assessment, in 2008, with the ITERS-R and ECERS-R in a nationally representative sample of child care centers. In addition to the ITERS-R and ECERS-R, which globally assess a broad range of aspects of the child care environment, we developed the Caregiver Interaction Profile (CIP) scales (see Helmerhorst, Riksen-Walraven, Vermeer, Fukkink, & Tavecchio, 2014) to gain a more detailed picture of what is generally considered the core aspect of child care quality (i.e., the quality of interactions between professional caregivers and the children in their care; Layzer & Goodson, 2006; Vandell & Wolfe, 2000). This is the first time the CIP scales have been applied in a nationally representative sample in combination with the ITERS-R and ECERS-R, yielding both a broad and detailed picture of the quality of care and education provided in Dutch centers.

Before specifying the aims of the study, we provide more information on the Dutch child care context. Today, the vast majority of Dutch mothers with young children (>70%) are employed, although mostly part time. Parents of children attending formal child care are generally highly educated, which is explained by the relatively high costs of formal care (Merens, Hartgers, & Van den Brakel, 2012). Children attend child care from age 3 months onward (when paid maternity leave ends) to age 4 years (when kindergarten starts). The vast majority of caregivers have characteristically received vocational training at the intermediate level (SPW-3), which is not specifically directed at very young children and focuses on care rather than education. This focus is also reflected in the typical daily program of child care centers in The Netherlands, which consists mostly of free play with little time allotted to structured educational activities (Organisation for Economic Co-operation and Development, 2000, 2006).

The first aim of the present study was to compare the quality of center-based child care in 2008 as assessed with the ITERS-R and ECERS-R to the quality found in earlier assessments, especially the previous assessment in 2005. The decrease in quality as assessed with the

ITERS-R and ECERS-R observed in The Netherlands between 1995 and 2005 was significant and substantial. In 1995, the average overall quality score on the ITERS/ECERS was 4.8 on a 7-point scale, which was relatively high in international perspective (Van IJzendoorn et al., 1998). In 2001, the average overall quality decreased to 4.3 (Gevers Deynoot-Schaub & Riksen-Walraven, 2005), and in 2005 it further decreased to 3.2, which is low, both in absolute terms and in international perspective (Vermeer et al., 2008). Two main factors have been proposed to explain the observed decrease in child care quality across the years 1995–2005. First, the explosive growth in the number of young children attending child care, together with a shortage of qualified staff and increased workload, may have put quality under pressure. A second possible explanation lies in changes in the vocational training of child care professionals in The Netherlands in the past decade. In earlier years, caregivers received specialized vocational training focused on working with very young children in a child care setting (LKC; Caregiver Child Care Centers), whereas from 1996 onward most caregivers have received a more general social pedagogical 3-year training (SPW-3; Social Pedagogical Work) that prepares them for working with a broader variety of age groups in different domains of care (Gevers Deynoot-Schaub & Riksen-Walraven, 2005; Vermeer et al., 2008).

It is difficult to predict whether child care quality has further declined since the last measurement in 2005 or whether the decline has stopped or even turned into an increase. The number of children attending child care centers has continued to increase since 2005 (in 2008, more than 31,000 children were on waiting lists for child centers; see Van Beem & Wever, 2008), so this probably unfavorable influence on quality remained. An important policy change that may have affected child care quality in The Netherlands in the past years is the introduction of the Child Care Act in 2005. This act redirected financial support from providers to parents to give parents more freedom to choose child care that fits their needs. This may lead to quality improvement if parents choose high-quality child care for their children, but it is questionable whether such an effect—if any (see also Plantenga, 2012; Vandell & Wolfe, 2000)—could already be visible 3 years after the introduction of the act. In addition, the introduction of the Child Care Act did not change regulations for quality monitoring, which could have led to a change in quality. In The Netherlands, quality monitoring is regulated by the Association of Community Health Services, which is responsible for (unannounced) inspection and enforcement of quality in child care centers. Based on this, there was no ground for formulating hypotheses regarding possible changes in the quality of child care since the former national quality assessment in 2005.

The second aim of the present study was to apply the newly developed CIP scales as an extension of the ITERS-R/ECERS-R in this large and nationally representative sample. The ITERS-R and ECERS-R are very useful in regular national quality assessments, because they provide a comprehensive picture of the quality of the child care environment in a broad set of domains, including caregiver–child interactions. In addition to this more global picture, we aimed to get a more in-depth picture of the interactions between professional caregivers and children by applying the CIP scales to video recordings made in the centers. The CIP scales were especially developed to be used in combination with the ITERS-R/ECERS-R in national quality assessments in Dutch child care centers. The CIP scales rate six important caregiver skills in interacting with 0- to 4-year-old children in a group setting (see the Method section). The results of a previous study supported the validity of the scales, suggesting that the CIP scales are indeed a valuable extension of the ITERS-R and ECERS-R (Helmerhorst et al., 2014).

The third aim of the study was to relate the process quality of child care, as measured by the ITERS-R/ECERS-R and the CIP scales, to three structural features of child care groups, namely, group size, caregiver–child ratio, and group type (age composition). Research has demonstrated the relation between process quality (especially caregiver–child interactions) and the structural features group size and caregiver–child ratio; in general, higher process quality is observed in smaller groups and with fewer children per caregiver (De Schipper, Riksen-Walraven, & Geurts, 2006; National Institute of Child Health and Human Development Early Child Care Research Network [NICHD ECCRN], 1996, 2002; Vandell & Wolfe, 2000). Therefore, we expected a smaller group size and fewer children per caregiver to be associated with higher process quality, especially with higher quality of caregiver–child interactions as measured with the CIP scales and, more globally, with the Interactions subscale of the ITERS-R/ECERS-R. With regard to age composition, we distinguished three types of groups: infant groups (0- to 2-year-olds), preschool groups (2- to 4-year-olds), and mixed-age groups (0- to 4-year-olds). The mixed-age group, with its relatively wide age range, is predominant in Dutch centers, whereas it is relatively unknown in other countries. Although there has been much discussion among professionals and parents in The Netherlands about the pros and cons of, especially, mixed-age versus infant groups both for children and for caregivers, relevant empirical evidence is still lacking. Therefore, we also explored the relation between group type and the process quality of care provided in the groups in the present representative sample.

Internationally, caregiver education is also considered an important structural quality feature, and research in various countries has shown it to be associated with child care quality (e.g., NICHD ECCRN, 2002; for a review, see Vandell & Wolfe, 2000). But in The Netherlands, earlier studies have consistently failed to show an association between caregiver education and child care quality. This is probably because of the lack of variation in education among caregivers in Dutch child care centers; in The Netherlands, the vast majority of caregivers have received the same vocational training at the intermediate level (see, e.g., De Schipper, Riksen-Walraven, & Geurts, 2007; Helmerhorst et al., 2014; Vermeer et al., 2008). Given this lack of variation in caregiver education we did not include it as a structural feature in the present study.

To summarize, this study had three aims. First, we examined the global process quality of child care in a nationally representative sample of child care centers in The Netherlands with the ITERS-R/ECERS-R and compared the results of that assessment to the previous assessment in 2005. Second, we used the newly developed CIP scales to observe the quality of caregiver–child interactions in more detail and related this quality to child care quality as assessed with the ITERS-R/ECERS-R. And third, we examined group size, caregiver–child ratio, and group type in relation to process quality, especially to the quality of caregiver–child interactions.

METHOD

Participants

A random sample of 200 child care centers was taken from a nationally representative pool of approximately 3,000 child care centers located across The Netherlands. The sample was stratified by region. Four backup samples per child care center were selected that could be approached in case the center did not participate. The participation rate was 33%. The main reason for refusal

was being too busy. Child care centers were approached by letter with an explanation of the purpose and procedures of the study and an invitation to cooperate. One week after the mailing, the center directors received a phone call to ask about their willingness to participate in the study. The final sample of 200 child care groups included 55 infant groups (0- to 2-year olds), 52 preschool groups (2- to 4-year olds), and 93 mixed-age groups (0- to 4-year olds). We randomly selected one care group per child care center. All caregivers of a selected care group were invited to participate in the study. The final sample include 425 caregivers, who were mostly female (98%), were born in The Netherlands (90%), had an average age of 32.7 years ($SD = 9.7$, range = 19–64), worked 27.9 hr a week ($SD = 6.5$) on average, and had an average of 8.1 years ($SD = 6.1$) of work experience in child care. The vast majority of the caregivers (81%) had completed the regular vocational education at the intermediate level. Only 3% had not (yet) completed this regular vocational education, and 12% had completed vocational education at a higher level. Just like the education at the intermediate level, the higher education was also more general and not specifically directed at early child care. This distribution shows that there was hardly any variation in caregiver education, which is in line with previous research in child care centers in The Netherlands (De Schipper et al., 2007; Helmerhorst et al., 2014; Vermeer et al., 2008). All participating caregivers and all parents of the children from the selected groups gave their active consent for filming procedures.

Procedure

Observations took place between September 2007 and January 2009. Each group was visited by a trained researcher from about 8 a.m. until approximately 3 p.m. The researcher completed the ITERS-R or ECERS-R and also made video recordings of caregivers in interactions with the children throughout the day (for applying the CIP scales afterward). Each individual caregiver was filmed for 8 to 10 min in four different situations: diapering, lunch/snack, free play, and transition between group activities. Caregivers were not given any specific instructions in advance and were asked to follow the regular program in order to capture a day as usual. In accordance with the ITERS-R/ECERS-R manual, the researcher had a short interview with one of the caregivers in the afternoon to obtain information on ITERS-R/ECERS-R items that could not be observed during the visit. Prior to the visit, caregivers had been sent questionnaires in order to gather individual background information (e.g., age, education, work experience, and work hours).

Measures

ITERS-R/ECERS-R. The ITERS-R (Harms et al., 2003) and the ECERS-R (Harms et al., 1998) were used to measure global process quality in the child care groups. The ITERS-R was developed for use in groups in which more than 50% of children are younger than 30 months, whereas the ECERS-R was developed for use in groups in which more than 50% of the children are between the ages of 30 and 48 months. We used these cutoff values to determine which of the two instruments was used in mixed-age groups (0- to 4-year-olds). The ITERS-R consists of 39 items and the ECERS-R consists of 43 items, and both constitute seven subscales: (a) Space and Furnishings, (b) Personal Care Routines, (c) Language, (d) Activities, (e) Interactions,

(f) Program Structure, and (g) Provisions for Parents and Staff. The Personal Care Routines and Provisions for Parents and Staff subscales were not used in the present study. Items are rated on a 7-point scale with descriptors for scores 1 (inadequate), 3 (minimal), 5 (good), and 7 (excellent). In addition to mean scores for the separate subscales, we also computed a total ITERS-R/ECERS-R mean score by averaging scores across all items. Based on the mean scores, groups were classified according to the quality levels *low* ($M < 3$), *moderate* ($3 \leq M < 5$) and *high* ($M \geq 5$).

Prior to data collection, eight observers were trained to use the ITERS-R, six observers were trained to use the ECERS-R, and three observers were trained to use both instruments. Each observer performed at least three onsite visits (range = 3–6) supervised by an expert trainer and followed by an item-by-item debriefing. Interobserver agreement of 80% (within 1 scale point) on three consecutive visits was required before observers were allowed to collect data independently. The average interobserver agreement was 88%, ranging from 81% to 95% for the different scales. Interobserver agreement checked during 10% of the visits was 89%.

The CIP scales. (For a more detailed description of development and first validation results, see Helmerhorst et al., 2014.) The CIP scales measure six caregiver interactive skills: (a) *sensitive responsiveness*, or the extent to which a caregiver recognizes children's individual emotional and physical needs and responds appropriately and promptly to their cues and signals; (b) *respect for autonomy*, or the extent to which a caregiver is nonintrusive but instead recognizes and respects the validity of children's intentions and perspectives; (c) *structuring and limit setting*, or the ability of a caregiver to clearly communicate expectations to children and structure a situation accordingly and to set clear and consistent limits for children's behavior; (d) *verbal communication*, or the frequency and quality of verbal interactions between the caregiver and children; (e) *developmental stimulation*, or the degree to which a caregiver deliberately attempts to foster children's development (e.g., motor development, cognitive development and creativity) while attuning the stimulation to the children's focus of attention, developmental level, and state; and (f) *fostering positive peer interactions*, or a caregiver's guidance of interactions between children in the child care group. Each of the six CIP scales is rated on a single 7-point Likert-type scale (7 = very high, 6 = high, 5 = moderate/high, 4 = moderate, 3 = moderate/low, 2 = low, 1 = very low). In the extensive description of each of the six scales, a general definition of the corresponding caregiver interactive skill is given first, followed by a brief description distinguishing scores at the high (6, 7), middle (3, 4, 5), and low (1, 2) ranges of the scale. Finally, for each of the scales, a detailed description of behavior for each of the seven scale points is provided. For a more comprehensive description of the CIP scales, see Helmerhorst et al. (2014).

Fifteen observers independently rated the videotaped caregiver–child interaction episodes. Observers attended six training sessions before applying the CIP scales. The training was followed by a test, and intraclass correlations were computed. Observers had to meet a criterion of .70 for each scale before they could start to observe. Observers independently rated the behavior of the caregiver on the six 7-point scales for each of four videotaped episodes. A mean score for each of the six skills was calculated per caregiver by averaging across the four episodes. For analyses at the level of the child care groups, scores of all caregivers per group were averaged. Analogous to the ITERS-R/ECERS-R, mean scores on the CIP scales were classified into three quality levels. The levels are labeled *inadequate* ($M < 3.5$), *moderate* ($3.5 \leq M < 4.5$), and *adequate to good* ($M \geq 4.5$).

Interrater reliability (i.e., intraclass correlations computed for 10% of the tapes) was .83, on average.

Group size and caregiver–child ratio. Group size was defined as the total number of children present in the group during the visit. Caregiver–child ratio was defined as the total number of children divided by the total number of caregivers in the group during the visit.

RESULTS

Global Process Quality Compared to Earlier Assessments

Table 1 presents the means and standard deviations for the ITERS-R/ECERS-R scores for the present assessment (2008) and the earlier assessments (1995, 2001, and 2005). The table shows that the declining trend for both the subscales and the total score from 1995 to 2001 and 2005 (see also Vermeer et al., 2008) continued to 2008. The *t* tests showed that the mean total score and the mean scores for the Space and Furnishings subscale and the Interactions subscale in the present assessment were substantially and significantly lower than the scores for the preceding assessment in 2005. Effect sizes were $d = 0.52$, $d = 0.50$, and $d = 1.01$, respectively. The effect size for the decrease in the total score was strong in terms of Cohen's (1988) definition ($d = 0.20$ is a weak effect, $d = 0.50$ a modest effect, and $d = 0.80$ a strong effect). In terms of the ITERS-R/ECERS-R scale point definitions, the present ITERS-R/ECERS-R total score ($M = 3.0$, $SD = 0.60$) was just at the minimal level. For the subscales, the lowest score was obtained for Activities ($M = 2.2$, $SD = 0.58$), followed by Space and Furnishings ($M = 3.0$, $SD = 0.78$), Language ($M = 3.4$, $SD = 1.08$), Program Structure ($M = 3.7$, $SD = 1.03$), and Interactions ($M = 3.8$, $SD = 1.17$).

TABLE 1
ITERS-R/ECERS-R Scores, Group Size, and Caregiver–Child Ratio in 1995, 2001, 2005, and 2008,
With *t* Values and Effect Sizes for Differences Between 2005 and 2008

Variable	1995 (N = 43)		2001 (N = 51)		2005 (N = 42)		2008 (N = 200)		2008 vs. 2005	
	M	SD	M	SD	M	SD	M	SD	<i>t</i>	Cohen's <i>d</i>
ITERS-R/ECERS-R										
Space and Furnishings	5.1	0.98	4.9	1.07	3.4	0.69	3.0	0.78	−3.08**	0.52
Personal Care Routines	4.3	1.06	4.4	1.11						
Language	4.7	1.01	3.7	1.30	3.5	1.13	3.4	1.08	−0.54	0.09
Activities	4.3	0.82	3.6	0.71	2.4	0.69	2.2	0.58	−1.96	0.33
Interactions	5.3	0.97	5.1	1.46	4.4	1.37	3.8	1.17	−2.93**	0.50
Program Structure	5.2	1.24			3.9	1.00	3.7	1.03	−1.15	0.20
Total ^a	4.8	0.61	4.3	0.74	3.6	0.57	3.0	0.60	−5.94**	1.01
Group size	11.1	4.0	9.1	2.6	9.7	2.6	10.9	3.2	2.28*	0.39
Caregiver–child ratio	0.27	0.12	0.26	0.08	0.22	0.05	0.21	0.05	−1.18	0.20

Note. ITERS-R/ECERS-R = Infant/Toddler Environment Rating Scale–Revised/Early Childhood Environment Rating Scale–Revised.

^aTotal scores for 1995 and 2001 are based on all items, including Personal Care Routines; total scores for 2005 and 2008 are based on all items, excluding Personal Care Routines.

* $p < .05$. ** $p < .01$.

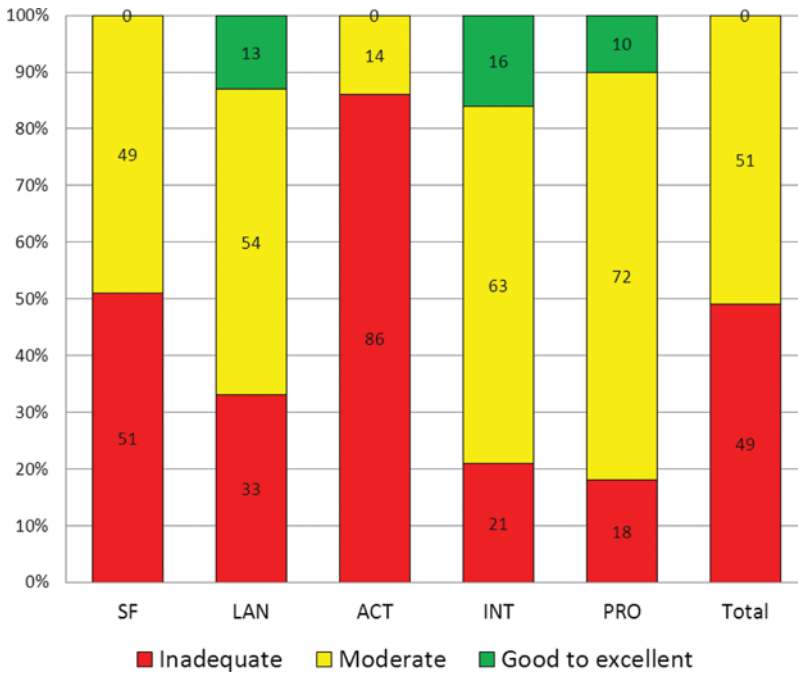


FIGURE 1 Percentages of care groups ($N = 200$) with mean scores of inadequate, moderate, and good to excellent on the Infant/Toddler Environment Rating Scale-Revised/Early Childhood Environment Rating Scale-Revised. SF = Space and Furnishings; LAN = Language; ACT = Activities; INT = Interactions; PRO = Program Structure.

Figure 1 shows the percentages of child care groups with high, moderate, and low scores in the present assessment in terms of the predefined ITERS-R/ECERS-R quality categories. With regard to the total score (see the rightmost bar in the figure), none of the groups scored in the high category, 51% scored moderate, and 49% scored low. In the preceding assessment in 2005 there were no high total scores either; 64% of the groups scored moderate and 36% scored low. This indicates that between 2005 and 2008, 13% of the groups shifted from the moderate to the low category.

As can be seen in Figure 1, results from the present assessment demonstrate that for the Space and Furnishings subscale and for the Activities subscale none of the groups scored high, and the vast majority of the groups scored low. For the Language, Interactions, and Program Structure subscales, a small percentage of groups (between 10% and 16%) scored high, the majority of the groups scored moderate, and between 18% and 33% of the groups scored low.

Caregiver Interactive Skills

The correlations among the CIP scales were all significant, ranging from .34 to .75. The means, standard deviations, and ranges for the CIP scores are presented in Table 2 (left-hand side). The table shows substantial variation among the means of the different CIP scales. The means for the

TABLE 2
Descriptive Statistics for the CIP Scales and Pearson Correlations Between the CIP Scales and
ITERS-R/ECERS-R Subscales

Variable	M	SD	Range	ITERS-R/ECERS-R (N = 200)					
				SF	LAN	ACT	INT	PRO	Total
CIP (N = 425)									
Sensitive responsiveness	4.81	0.93	1.50–7.00	.24**	.34**	.29**	.40**	.36**	.45**
Respect for autonomy	4.51	0.84	1.25–6.75	.28**	.38**	.38**	.36**	.34**	.49**
Structuring and limit setting	4.90	1.17	2.00–7.00	.11	.18*	.22**	.18*	.17*	.24**
Verbal communication	3.60	0.90	1.25–6.00	.28**	.28**	.28**	.33**	.26**	.42**
Developmental stimulation	2.16	0.93	1.00–5.50	.22**	.25**	.23**	.24**	.24**	.33**
Fostering positive peer interactions	1.72	0.83	1.00–7.00	.13	.39**	.27**	.24**	.23**	.35**

Note. Correlations between all CIP scales and ITERS-R/ECERS-R subscales are at the care group level (N = 200). CIP = Caregiver Interaction Profile; ITERS-R/ECERS-R = Infant/Toddler Environment Rating Scale–Revised/Early Childhood Environment Rating Scale–Revised; SF = Space and Furnishings; LAN = Language; ACT = Activities; INT = Interactions; PRO = Program Structure.

* $p < .05$ (two-tailed). ** $p < .01$ (two-tailed).

most basic caregiver skills (sensitive responsiveness, respect for autonomy, and structuring and limit setting) were in the adequate to good category (score ≥ 4.5), whereas the means for the more educational skills were moderate ($3.5 \leq \text{score} < 4.5$) for verbal communication and even inadequate (score < 3.5) for developmental stimulation and fostering positive peer interactions.

Figure 2 presents the percentages of groups that scored adequate to good, moderate, and inadequate for the six CIP caregiver skills. The vast majority of the caregivers scored adequate to good on sensitive responsiveness, respect for autonomy, and structuring and limit setting. Less than 10% of the caregivers scored inadequate on these scales. For the remaining skills we found a quite different pattern. Here only a tiny minority of the caregivers had a high score (13%, 3%, and 1%, respectively, for verbal communication, developmental stimulation, and fostering positive peer interactions), whereas almost half to almost all caregivers had inadequate scores (42%, 89%, and 99%, respectively, for verbal communication, developmental stimulation, and fostering positive peer interactions). To summarize, the vast majority of caregivers scored adequate to good on the more basic caregiving skills sensitive responsiveness, respect for autonomy, and structuring and limit setting, whereas the vast majority scored inadequate for the more educational skills developmental stimulation and fostering positive peer interactions.

The right side of Table 2 shows the correlations between the CIP scores and the ITERS-R/ECERS-R scores. In general, the correlations were significant and moderate (ranging from .11 to .49), indicating that the CIP scales and the ITERS-R/ECERS-R, as intended, measure distinct but related aspects of the same construct (i.e., process quality).

Process Quality as Related to Group Size, Caregiver–Child Ratio, and Group Type

Table 1 shows the means and standard deviations for group size and caregiver–child ratio. Compared to 2005, the average group size increased significantly in 2008 by more than one

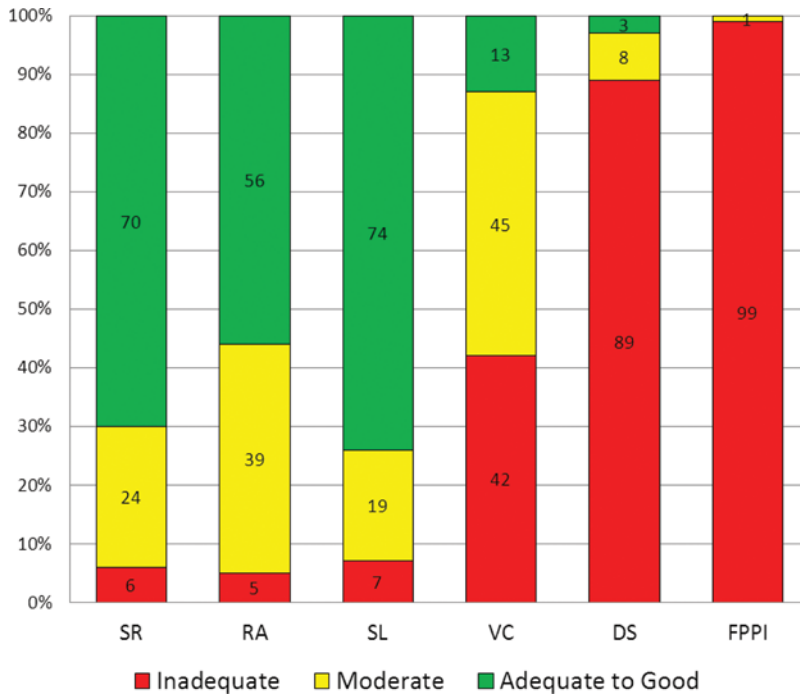


FIGURE 2 Percentages of care groups ($N = 199$) with mean scores of inadequate, moderate, and adequate to good on the Caregiver Interaction Profile scales. SR = sensitive responsiveness; RA = respect for autonomy; SL = structuring and limit setting; VC = verbal communication; DS = developmental stimulation; FPPI = fostering positive peer interactions.

child per group to 10.9 ($SD = 3.2$). With regard to caregiver–child ratio, the difference from 2005 was not significant. On average, there was one caregiver per five children ($M = 0.21$, $SD = 0.05$).

Correlations between group size and caregiver–child ratio on the one hand and the ITERS-R/ECERS-R subscales on the other were significant for the Interactions subscale ($-.18$, $p < .05$; and $.18$, $p < .01$, respectively), indicating that the quality of interactions was lower with larger group sizes and more children per caregiver. The other subscales were not significantly related to group size or caregiver–child ratio. Group size was not significantly related to any of the CIP scales. Caregiver–child ratio, however, was significantly related to two of the scales, that is, verbal communication ($-.22$, $p < .01$) and developmental stimulation ($-.15$, $p < .05$), indicating, surprisingly, that caregivers scored higher on verbal communication and developmental stimulation in groups with more children per caregiver.

Finally, we examined whether process quality, as measured with the ITERS-R/ECERS-R and the CIP scales, was related to group type (infant, preschool, or mixed age). Table 3 shows means and standard deviations for the ITERS-R/ECERS-R subscales and the CIP scales separately for the three group types. For the ITERS-R/ECERS-R, we conducted a multivariate analysis of variance with type of group as a between-subjects factor and the ITERS-R/ECERS-R subscales as dependent variables. We found a significant overall effect of group type on the

TABLE 3
Means and Standard Deviations for the Three Different Group Types

Variable	Infant (I) (n = 55)		Preschool (P) (n = 52)		Mixed Age (M) (n = 93) ^a		F	Significant Contrasts ^b
	M	SD	M	SD	M	SD		
CIP								
Sensitive responsiveness	4.57	0.90	5.05	0.65	4.85	0.75	5.21**	I < P
Respect for autonomy	4.26	0.77	4.73	0.60	4.55	0.68	6.53**	I < P, M
Structuring and limit setting	4.96	1.00	5.12	0.66	4.81	1.04	1.78	
Verbal communication	3.23	0.66	3.87	0.64	3.72	0.73	13.21**	I < P, M
Developmental stimulation	1.87	0.68	2.34	0.80	2.30	0.86	6.38**	I < P, M
Fostering positive peer interactions	1.49	0.53	1.82	0.65	1.79	0.76	4.15*	I < P, M
ITERS-R/ECERS-R								
Space and Furnishings	2.70	0.63	3.29	0.68	3.00	0.62	7.58**	I < P
Language	3.45	1.09	3.56	1.08	3.28	1.08	2.29	
Activities	2.31	0.56	2.13	0.50	2.27	0.63	0.58	
Interactions	3.81	1.25	3.70	0.97	3.75	1.25	1.65	
Program Structure	3.47	1.18	3.77	0.93	3.71	0.98	2.32	
Total	2.93	0.60	3.08	0.60	3.01	0.61		

Note. CIP = Caregiver Interaction Profile; ITERS-R/ECERS-R = Infant/Toddler Environment Rating Scale-Revised/Early Childhood Environment Rating Scale-Revised.

^aFor one mixed-age group, CIP scores were not available because of a technical problem with the camera. ^bScheffé post hoc test.

* $p < .05$. ** $p < .01$.

ITERS-R/ECERS-R, $F(12, 384) = 3.95$, $p < .00$, Wilks's $\Lambda = 0.792$. As can be seen in Table 3, subsequent univariate analyses revealed a significant effect of group type only for the Space and Furnishings subscale. Post hoc tests showed that scores for Space and Furnishings were significantly lower in infant groups compared to the preschool groups. For the CIP scales, we conducted a multivariate analysis of variance with type of group as a between-subjects factor and the six CIP scales as dependent variables. Results demonstrated a significant overall effect of group type on the CIP scales, $F(12, 381) = 1.87$, $p < .01$, Wilks's $\Lambda = 2.394$. Subsequent analyses of variance showed significant effects of group type for all CIP scales except for structuring and limit setting (see Table 3). Post hoc tests showed that caregivers in infant groups scored significantly lower on sensitive responsiveness than caregivers in preschool groups. For respect for autonomy, verbal communication, developmental stimulation, and fostering positive peer interactions, caregivers in infant groups also scored significantly lower than caregivers in the mixed-age groups.

DISCUSSION

This study shows, first of all, that the decreasing trend in the quality of child care in Dutch centers as measured with the ITERS-R/ECERS-R across the years 1995–2005 continued in the 3 years thereafter. The ITERS-R/ECERS-R total score was significantly and substantially lower in 2008 than in 2005, with almost half of the groups now scoring below the minimal level.

The CIP scales, which were used for the first time in a nationally representative sample of child care centers, provided a more detailed picture of the core aspect of process quality, namely, the caregiver–child interaction. The CIP scores revealed the typical strengths and weaknesses of caregivers' interactive skills in Dutch child care centers, with relatively high scores for the more basic caregiving skills sensitive responsiveness, respect for autonomy, and structuring and limit setting and clearly lower scores for the more educational skills verbal communication, developmental stimulation, and fostering positive peer interactions.

The research design of this study does not allow for drawing definite conclusions about possible causes for the further decline in global process quality. As outlined in the introduction, the number of children entering child care has continued to increase since the former quality assessment, which may partly explain the decrease in quality. Another possible explanation for the further decline might lie in changes in caregiver vocational training. As also mentioned in the introduction, in 1996 caregiver vocational training with a specific focus on working with young children (LKC) changed to a more general social pedagogical training (SPW-3). Between 2005 and 2008, however, caregiver education remained unchanged. Therefore, caregiver vocational training is not a plausible explanation for the decline in quality across those years.

Whether the introduction of the Child Care Act in 2005—a major policy change since the former quality assessment—contributed to the further decline in child care quality between 2005 and 2008 also cannot be stated with certainty. As noted in the introduction, the Child Care Act introduced a market-driven system that could lead to quality improvement if it stimulates parents to choose more high-quality child care for their children. The continuing decrease in quality after the introduction of the act suggests that the child care market did not (yet) work as intended. This apparent market failure may have different causes, the most important of being a lack of information related to pedagogical quality for parents (Vandell & Wolfe, 2000). In The Netherlands, information regarding the comparative quality of care provided in child care centers is not publicly available. In addition, it appears to be difficult for parents to reliably judge the quality of child care themselves based on their own observations; they rate the quality of Dutch child care centers as more than sufficient (Berden & Kok, 2009; Kok, Groot, Mulder, Sadirai, & Van Ham, 2005), whereas the results of the present study point in a different direction. In choosing a child care center for their children, Dutch parents are primarily led by convenience factors like availability (length of waiting lists) and distance to home and/or work (Berden & Kok, 2009; Kok et al., 2005). And parents who are not satisfied with the quality of the care provided in a center hardly ever switch to another center, even when alternative options are available (Berden & Kok, 2009), because of the emotional burden on the children (Plantenga, 2012). The present findings help explain why the child care market that was introduced by the Child Care Act in 2005 evidently did not lead to the intended improvement in the quality of child care in subsequent years.

This study is the first to have applied the CIP scales in a nationally representative sample of child care centers together with the ITERS-R/ECERS-R to get more detailed information about the quality of caregiver–child interactions. The results of the present study underscore those of a previous study with a smaller and nonrepresentative sample (Helmerhorst et al., 2014), indicating that the CIP scales are indeed a valuable extension of the ITERS/ECERS. The average score on the ITERS-R/ECERS-R Interactions subscale was 3.8 on a 7-point scale (moderate), which is in line with the moderate average score across the six CIP scales ($M = 3.6$ on a 7-point scale). But the results from the CIP scales also demonstrate substantial differences among the different caregiver interactive skills, with adequate to good scores for the more basic caregiving skills of

sensitive responsiveness ($M=4.8$), respect for autonomy ($M=4.5$), and structuring and limit setting ($M=4.9$) to inadequate scores for the more educational skills of verbal communication ($M=3.6$), developmental stimulation ($M=2.2$), and fostering positive peer interactions ($M=1.7$). The CIP scores also revealed interesting differences in caregiver–child interactions in different group types (see Table 3) that were not visible in the ITERS-R/ECERS-R scores. Overall, these results underscore the fact that the CIP scales provide valuable additional information on caregiver–child interactions next to the ITERS-R/ECERS-R Interactions subscale.

Our present finding that caregivers scored higher on the more basic interactive skills than on the more educational skills of the CIP is in line with findings in earlier studies with smaller and nonrepresentative samples of caregivers in Dutch centers (Albers, Riksen-Walraven, & de Weerth, 2010; Helmerhorst et al., 2014) and with findings with the Classroom Assessment Scoring System in the United States (Guo, Piasta, Justice, & Kaderavek, 2010; La Paro, Pianta, & Stuhlman, 2004; LoCasale-Crouch et al., 2007; Mashburn et al., 2008; Thomason & La Paro, 2009). The lower scores for the more educational skills can also be explained by the focus on care rather than education in Dutch child care centers (Helmerhorst et al., 2014; Organisation for Economic Co-operation and Development, 2006). The positive influence of caregivers' educational or instructional support on children's cognitive and language development, however, has been demonstrated in numerous studies (e.g., Albers et al., 2010; Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Howes et al., 2008; Mashburn et al., 2008). These results endorse the importance of the more educational interactive skills and encourage policies and practices to improve the quality of caregivers' educational skills through training and professional development (LoCasale-Crouch et al., 2007; Mashburn et al., 2008).

In a previous study, in which we introduced the CIP scales and examined their validity, we were unable to obtain reliable ratings for structuring and limit setting, because the relevant caregiver behaviors occurred too infrequently on the videotaped interaction episodes. We assumed this to be due to the filming procedure, which was based on a fixed preset time schedule across the day. This yielded a large proportion of video fragments in free play situations, during which caregivers mostly did not have an active structuring role toward the children. Therefore, we adapted the filming procedure to film caregivers in four situations: diapering, lunch/snack, free play, and a transition between group activities. The transition moment was specifically chosen to rate the skill structuring and limit setting. With the adapted filming procedure used in the present study the structuring and limit setting scale could be reliably applied. The scores for the other CIP scales were comparable to those in the former study (Helmerhorst et al., 2014), suggesting that the adaptation of the filming procedure did not affect the CIP results.

Our expectation that a smaller group size and fewer children per caregiver would be associated with higher quality caregiver–child interactions was supported for the ITERS-R/ECERS-R Interactions subscale but not for the CIP scales. Group size was unrelated to any of the CIP scales, and the association between the caregiver–child ratio and two of the CIP scales (i.e., verbal communication and developmental stimulation) was even in contrast to our expectation: Caregivers scored significantly higher instead of lower on the CIP scales verbal communication and developmental stimulation with more children per caregiver. With regard to the lack of association between group size and CIP scores, this lack of association could be characteristic for The Netherlands, given that this result is in line with findings of earlier Dutch child care studies. Van IJzendoorn et al. (1998) and Vermeer et al. (2008) found group size to be unrelated to caregivers' scores on the Caregiver Interaction Scale (Arnett, 1989), and De Schipper et al.

(2007) found it to be unrelated to caregiver–child interactions as observed with the Caregiver Interaction Scales as well as with rating scales that are comparable to the CIP scales. A methodological explanation could be that the lack of correlation is due to a restricted range of group size in Dutch centers, but the score distribution for group size—just as that for caregiver–child ratio—is not substantially different from the distribution reported in other countries (see Vandell & Wolfe, 2000). This raises the question of whether any other specific characteristics of the Dutch situation might explain the lack of association between group size and caregiver–child interactions. What comes to mind is that the parents of children attending child care are relatively highly educated, which may contribute to a smaller number of children per group with problem behaviors. Groups with fewer highly demanding children could be easier for caregivers to manage, which could explain why a few more children per group do not go at the cost of the quality of caregiver–child interactions. Of course this is merely an assumption, and more research is warranted to examine this further.

The unexpected finding that more children per caregiver was related to higher caregiver scores on the CIP verbal communication and developmental stimulation scales is quite puzzling. A possible explanation for the higher verbal communication scores for caregivers who have to interact with more children could be that more children per caregiver increases the number of children asking for attention in the group, which may lead to more verbal communication between caregiver and children—not only in caregivers’ interactions with individual children but also in group-directed verbal interactions. The significantly higher levels of developmental stimulation provided by caregivers who have to care for more children remain puzzling. It cannot be ruled out that more highly skilled caregivers are placed in groups with more children per caregiver; more research is needed to examine this. Moreover, given that the unexpected negative correlations between caregiver–child ratio and the two CIP scales were quite low and applied to only two of the six CIP scales, replication is needed before definitive conclusions can be drawn about the relation between CIP scores and caregiver–child ratio.

When comparing the quality of caregiver–child interactions in groups with different age compositions, we found that scores on the CIP sensitive responsiveness scale were significantly lower in infant groups than in preschool groups. Furthermore, caregiver scores for respect for autonomy, verbal communication, developmental stimulation, and fostering positive peer interactions were significantly lower in infant groups than in preschool groups and mixed-age groups. This is in line with the results of an earlier study in Dutch child care centers that also showed that having more children younger than the age of 2 in a care group significantly predicted lower quality caregiver–child interactions (De Schipper et al., 2007). Given that sensitive responsiveness is generally considered to be the most basic aspect of caregiver behavior in interaction with infants, the significant lower score for sensitive responsiveness in infant groups is a cause for concern and needs attention.

Limitations

The question of whether individual infants are better off in mixed-age groups than in infant groups cannot be answered in this study, because the CIP scales were designed to reflect caregivers’ behavior toward children in the care group in general and not their interactions with individual children. Comparing the quality of care for infants in infant groups versus mixed-age

groups requires observations and rating scales that focus on the children instead of the caregivers during caregiver–child interactions (e.g., the Observational Record of the Caregiving Environment; see NICHD ECCRN, 1996).

The response rate of the child care centers in the present study (33%) was comparable to that in our former study using the CIP scales (27%; Helmerhorst et al., 2014) but somewhat lower than in the former national quality assessment (49%; Vermeer et al., 2008). It is unclear why the response rate in the present study was lower. The procedures for recruitment in the present study were exactly the same as in 2005, and the main reason for refusal to participate in the present study (i.e., being too busy) was also the same as in the 2005 assessment. Unfortunately, it is not possible to draw any conclusions about the causes of this lower response rate based on the present data.

Although the present study sample seemed to be representative for The Netherlands, the results of this study cannot be automatically generalized to other countries given the typical characteristics of the Dutch child care context, outlined in the introduction to this article. Furthermore, it should be noted that the CIP scales were developed for use in child care centers in The Netherlands and may therefore reflect Dutch values and other characteristics of the Dutch child care context (see Helmerhorst et al., 2014). This should be taken into account when applying the scales in other countries.

However, although the results may be typical for the Dutch context, they may also be interesting for an international audience, because repeated child care quality assessments are quite rare in Europe and may contribute to understanding how quality is affected by changing circumstances and policies in a different context. The study may increase awareness that the quality of child care should always be considered in its own national and cultural context. In fact, the Dutch research program of which the present study is part (including repeated quality assessments and the development of culturally valid assessment instruments like the CIP scales) started because there were doubts about whether the results of an NICHD study in the United States (NICHD ECCRN, 2002) could be generalized to the Dutch context.

Implications for Policy and Practice

The continuing and substantial decrease in the quality of child care in The Netherlands as measured with the ITERS-R/ECERS-R in the 3 years since the former national quality assessment shows that the quality of child care is far from stable and that regular assessments are therefore needed to monitor child care quality.

The present study is the first to have used the CIP scales together with the ITERS-R/ECERS-R in assessing the quality of child care in a nationally representative sample of child care centers. The results underscore that the CIP scales provide valuable additional information beyond the ITERS-R/ECERS-R, and it is therefore recommended that the CIP scales be included in future quality assessments.

The low average scores on the more educational versus the more basic CIP scales indicate that the former skills (i.e., verbal communication, developmental stimulation, and fostering positive peer interactions) should get more attention in caregiver education and training in general. The large variability in CIP scores among caregivers, however, suggests that an individual approach in training may also be useful. This is underscored by the moderate intercorrelations between the CIP scales, which demonstrate that a high score for one skill does not necessarily

imply a high score for other skills. This suggests that training programs should be well attuned to the interaction skill profile of individual caregivers. We recently developed a video-feedback training program that aims to improve professional caregivers' interactive skills based on their individual profile of scores on the CIP scales. The effects of this training are presently being examined and will be reported in the near future.

REFERENCES

- Albers, E. M., Riksen-Walraven, J. M. A., & de Weerth, C. (2010). Developmental stimulation in child care centers contributes to young infants' cognitive development. *Infant Behavior & Development, 33*, 401–408. doi:10.1016/j.infbeh.2010.04.004
- Arnett, J. (1989). Caregivers in day-care centers: Does training matter? *Journal of Applied Developmental Psychology, 10*, 541–552. doi:10.1016/0193-3973(89)90026-9
- Belsky, J., Vandell, D. L., Burchinal, M., Clarke-Stewart, K. A., McCartney, K., Owen, M. T., & the NICHD Early Child Care Research Network. (2007). Are there long-term effects of early child care? *Child Development, 78*, 681–701. doi:10.1111/j.1467-8624.2007.01021.x
- Berden, C., & Kok, L. (2009). *Ontwikkelingen op de markt voor kinderopvang 2004–2008* [Developments for the market in child care 2004–2008]. Amsterdam, The Netherlands: Stichting Economisch Onderzoek.
- Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child care outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly, 25*, 166–167. doi:10.1016/j.ecresq.2009.10.004
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- De Schipper, E. J., Riksen-Walraven, J. M. A., & Geurts, S. A. E. (2006). Effects of child-caregiver ratio on the interactions between caregivers and children in child-care centers: An experimental study. *Child Development, 77*, 861–874. doi:10.1111/j.1467-8624.2006.00907.x
- De Schipper, E. J., Riksen-Walraven, J. M. A., & Geurts, S. A. E. (2007). Multiple determinants of caregiver behavior in child-care centers. *Early Childhood Research Quarterly, 22*, 312–326. doi:10.1016/j.ecresq.2007.04.004
- Gevers Deynoot-Schaub, M. J. J. M., & Riksen-Walraven, J. M. A. (2005). Child care under pressure: The quality of Dutch center in 1995 and 2001. *Journal of Genetic Psychology, 166*, 280–296. doi:10.3200/GNTP.166.3.280-296
- Guo, Y., Pianta, S. B., Justice, L. M., & Kaderavek, J. N. (2010). Relations among preschool teachers' self-efficacy, classroom quality, and children's language and literacy gains. *Teaching and Teacher Education, 26*, 1094–1103. doi:10.1016/j.tate.2009.11.005
- Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early Childhood Environment Rating Scale–Revised*. New York, NY: Teachers College Press.
- Harms, T., Cryer, D., & Clifford, R. M. (2003). *Infant/Toddler Environment Rating Scale–Revised*. New York, NY: Teachers College Press.
- Helmerhorst, K. O. W., Riksen-Walraven, J. M. A., Vermeer, H. J., Fukkink, R. G., & Tavecchio, L. W. C. (2014). Measuring interactive skills of caregivers in child care centers: Development and validation of the Caregiver Interaction Profile scales. *Early Education & Development, 25*, 770–790. doi:10.1080/10409289.2014.840482
- Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Early Childhood Research Quarterly, 23*, 27–50. doi:10.1016/j.ecresq.2007.05.002
- Kok, L., Groot, I., Mulder, J., Sadirai, K., & Van Ham, M. (2005). *De markt voor kinderopvang in 2004* [The child care market in 2004]. Amsterdam, The Netherlands: Stichting Economisch Onderzoek.
- La Paro, K. M., Pianta, R., & Stuhlman, M. (2004). The Classroom Assessment Scoring System: Findings from the prekindergarten year. *The Elementary School Journal, 104*, 409–426.
- Layzer, J. L., & Goodson, B. D. (2006). The "quality" of early care and education settings. *Evaluation Review, 30*, 556–576. doi:10.1177/0193841X06291524
- LoCasale-Crouch, J., Konold, T., Pianta, R., Howes, C., Burchinal, M., Bryant, D., . . . Barbarin, O. (2007). Observed classroom quality profiles in state-funded pre-kindergarten programs and associations with teacher, program, and classroom characteristics. *Early Childhood Research Quarterly, 22*, 3–17. doi:10.1016/j.ecresq.2006.05.001

- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., . . . Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development, 79*, 732–749. doi:10.1111/j.1467-8624.2008.01154.x
- Merens, A., Hartgers, M., & Van den Brakel, M. (2012). *Emancipatiemonitor 2012* [Emancipation Monitor 2012]. The Hague, The Netherlands: Sociaal en Cultureel Planbureau/Centraal Bureau voor de Statistiek.
- National Institute of Child Health and Human Development Early Child Care Research Network. (1996). Characteristics of infant child care: Factors contributing to positive caregiving. *Early Childhood Research Quarterly, 11*, 269–306. doi:10.1016/S0885-2006(96)90009-5
- National Institute of Child Health and Human Development Early Child Care Research Network. (2002). Child-care structure → process → outcome: Direct and indirect effects of child-care quality on young children's development. *Psychological Science, 13*(3), 199–206.
- Organisation for Economic Co-operation and Development. (2000). *Early childhood education and care policy in The Netherlands. Background report to the OECD-project*. The Hague, The Netherlands: Ministry of Health, Welfare & Sport, Ministry of Education, Culture and Science.
- Organisation for Economic Co-operation and Development. (2006). *Starting Strong II: Early childhood education and care*. Paris, France: Author.
- Plantenga, J. (2012). Local providers and loyal parents: Competition and consumer choice in the Dutch childcare market. In L. Lloyd & H. Penn (Eds.), *Childcare markets: Can they deliver an equitable service?* (pp. 63–78). Bristol, England: Policy Press. doi:10.1332/policypress/9781847429339.003.0004
- Thomason, A. C., & La Paro, K. M. (2009). Measuring the quality of teacher-child interactions in toddler child care. *Early Education & Development, 20*, 285–304. doi:10.1080/10409280902773351
- Tietze, W., Cryer, J., Bairrão, D., Palacios, J., & Wetzel, G. (1996). Comparisons of observed process quality in early child care and education programs in five countries. *Early Childhood Research Quarterly, 11*, 447–475. doi:10.1016/S0885-2006(96)90017-4
- Van Beem, M., & Wever, Y. (2008). *Wachttijsten en wachttijden buitenschoolse opvang en dagopvang (2^e meting)* [Waiting lists and waiting periods in after school care and child daycare centers (2nd measurement)]. Den Haag, The Netherlands: B&A Consulting bv.
- Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., Vandergrift, N., & NICHD Early Child Care Research Network. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD Study of Early Child Care and Youth Development. *Child Development, 81*, 737–756. doi:10.1111/j.1467-8624.2010.01431.x
- Vandell, D. L., & Wolfe, B. (2000). *Child care quality: Does it matter and does it need to be improved?* Madison: University of Wisconsin–Madison, Institute for Research on Poverty.
- Van IJzendoorn, M. H., Tavecchio, L. W. C., Stams, G. J. J. M., Verhoeven, M. J. E., & Reiling, E. J. (1998). Quality of center day care and attunement between parents and caregivers: Center day care in cross-national perspective. *The Journal of Genetic Psychology, 159*, 437–454.
- Vermeer, H. J., Van IJzendoorn, M. H., De Kruif, R. E. L., Fukkink, R. G., Tavecchio, L. W. C., Riksen-Walraven, J. M. A., & Van Zeijl, J. (2008). Child care in The Netherlands: Trends in quality over the years 1995–2005. *Journal of Genetic Psychology, 169*, 360–385.