



## UvA-DARE (Digital Academic Repository)

### Child care quality in the Netherlands: From quality assessment to intervention

Helmerhorst, K.O.W.

**Publication date**

2014

**Document Version**

Final published version

[Link to publication](#)

**Citation for published version (APA):**

Helmerhorst, K. O. W. (2014). *Child care quality in the Netherlands: From quality assessment to intervention*. [Thesis, fully internal, Universiteit van Amsterdam].

**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, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

**CHAPTER 6**  
**Summary, Conclusions and General**  
**Discussion**



Foto: Joods Historisch Museum



## 6.1 Summary of the research project

The studies presented in this thesis have been conducted as part of the Netherlands Consortium for Research in Child Care (Nederlands Consortium Kinderopvang Onderzoek, NCKO) research program, which has received funding from the Dutch Ministry of Social Affairs and Employment since 2002. The general purpose of the NCKO project was to assess *and* to improve child care quality for 0- to 4-year-olds in The Netherlands. The main aim of this thesis has been the development and evaluation of an intervention program to improve the quality of child care (study 3 and 4).

The first study describes the validation of the Caregiver Interaction Profile (CIP) scales. Next, this thesis reports on the application of the CIP scales in combination with the widely used Infant/Toddler Environment Rating Scale-Revised (ITERS-R; Harms, Cryer, & Clifford, 2003) and Early Childhood Environment Rating Scale-Revised (ECERS-R; Harms, Clifford, & Cryer, 1998) in a nationally representative sample of Dutch child care centers. Outcomes of this second study were the starting point for quality improvement. To this purpose, we developed a broad intervention program that aimed to improve both the quality of caregiver-child interactions as assessed with the CIP scales (described in study 3) and the more global quality of the child care environment as measured with the ITERS-R/ECERS-R (described in study 4) in the same care group.

**Study 1.** The first study, presented in Chapter 2, describes the development and validation of the CIP scales, which rate six caregiver behaviors that are assumed to play an important role in fostering the wellbeing and development of young children from birth onwards: *sensitive responsiveness, respect for autonomy, structuring and limit setting, verbal communication, developmental stimulation and fostering positive peer interactions.*

We examined *convergent validity* with another, somewhat comparable observational rating scale for caregiver interactive behavior and with a measure for global process quality. *Discriminant validity* was examined using ratings of caregivers' temperamental sociability, and *predictive validity* was examined by relating the CIP scales with children's competence and behavior problems as rated by their caregivers. Child care groups were visited by two trained researchers. The first researcher observed overall quality of the child care environment with the ITERS-R (Harms et al., 2003) and ECERS-R (Harms et al., 1998), while the second researcher recorded caregiver-child interactions on videotape in 8-10 minute cycles based on a fixed time schedule for later observation with the CIP scales. During a second visit three months later, caregiver-child interactions were recorded once more to determine test-retest reliability for the CIP scales and caregivers completed a questionnaire to rate children's competence and behavior problems.

Results showed promising preliminary evidence supporting the reliability and convergent, discriminant, and predictive validity of the CIP scales. The structuring

and limit setting scale could not be applied in this study because the relevant caregiver behaviors occurred too infrequently during the videotaped episodes (therefore, we adapted the video recording procedure for future studies – see study 2 below). Results of a multilevel analysis demonstrated that the CIP scales provide a unique interaction profile of individual caregiver skills; variation of caregiver behavior was for the most part situated at individual caregiver level, with relatively little variance at group and center level.

**Study 2.** The second study (Chapter 3) reports on the results of the second NCKO quality assessment, in 2008, in a nationally representative sample of 200 child care centers for 0- to 4-year-olds in The Netherlands. Global quality of the child care environment as measured with the ITERS-R and ECERS-R was compared with global quality at the prior assessment in a nationally representative sample in 2005. The newly developed CIP scales were applied for the first time to assess caregiver interactive skills in a nationally representative sample. Furthermore, child care quality was examined in relation to three structural features of child care groups, namely, group size, caregiver-child ratio, and group type (age composition).

Care groups were visited by a researcher who completed the ITERS-R/ECERS-R and made video recordings of each caregiver in interaction with children in four different situations, i.e., diapering, lunch/snack, free play, and transitions (8-10 minutes per situation). Video recordings were rated afterwards with the CIP scales by an independent and trained coder.

Results showed a significant decline in global quality as assessed with ITERS-R/ECERS-R scores in 2008 as compared to 2005, with almost half of the groups scoring below the minimal level. Outcomes on the CIP scales revealed relatively high scores for the more basic caregiver 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. 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). Group size was not related to any of the CIP scales. In contrast to our expectations, the significant relation between caregiver-child ratio and two of the CIP scales, i.e., verbal communication and developmental stimulation, indicated that caregivers scored higher on verbal communication and developmental stimulation in groups with more children per caregiver.

The low CIP scores on the more educational versus the more basic caregiver interactive skills indicate that these educational skills deserve extra attention in caregiver education and training.

**Study 3.** This study, presented in Chapter 4, examined the effects of a video feedback training for caregivers that aimed to strengthen quality of the caregiver-child interactions as assessed with the CIP scales. The training was based on the conceptual framework underlying the CIP scales and aimed to improve the six CIP skills.

The study sample included 68 mixed-age groups from 33 child care centers, with 35 groups randomly assigned to the experimental condition and 33 groups to the control condition. A total of 139 caregivers participated in the study. To assess the effects of the training, the quality of caregiver-child interactions was assessed with the CIP scales at pretest, posttest, and follow-up three months after the posttest. Following the procedure described in study 2, each caregiver was filmed in four situations (diapering, lunch/snack, structured play, and transition) for 8-10 minutes per situation. The ITERS-R/ECERS-R were also administered in the groups; these measures were used to control for possible effects of the parallel part of the intervention program (described in study 4).

Two weeks after the pretest, the video-feedback trainer visited the caregivers at the child care center for the first time for the video-feedback training, which consisted of five weekly meetings in total. Per group, all participating caregivers were trained individually with feedback on their own videotaped interaction episodes in the first four sessions. The fifth and final session was used to share experiences and video observations with the other trained caregiver from the same group.

Results showed a significant positive training effect on all six CIP skills at posttest. Three months after the posttest, at follow-up, caregivers who had received training still scored significantly higher on four of the six CIP skills: sensitive responsiveness, respect for autonomy, verbal communication, and fostering positive peer interactions. Although training effects were relatively small in absolute terms, they appear meaningful in terms of decreasing percentages of caregivers with inadequate scores and increasing percentages of caregivers scoring in the adequate to good range.

The results demonstrate that it is possible to enhance the quality of caregiver-child interactions through this video feedback training based on the CIP scales. Extending the CIP training with additional booster sessions is suggested to further improve the effectiveness of the training. Possible ways to implement the training in practice and education are discussed.

**Study 4.** The fourth study, presented in Chapter 5, examined the effects of the second part of the broader intervention program, namely an on-site consultancy program directed at center directors that aimed to improve the quality of the child care environment as reflected in the ITERS-R/ECERS-R (i.e., space, furnishings, materials, and program structure). The consultancy program ran parallel to the video-feedback training for caregivers in the same care group.

Participating groups were the same as in study 3, i.e., 68 mixed-age groups from 33 child care centers, with 35 groups randomly assigned to the experimental condi-

tion and 33 groups to the control condition. A total of 14 center directors participated in the consultancy program with one group, 6 center directors with two groups, and 3 center directors with three groups. All groups were visited for pretest, posttest, and follow-up three months after the posttest to complete the ITERS-R and ECERS-R that were used to examine the effects of the consultancy program. CIP scores for the caregivers in the groups, obtained at pretest, posttest, and follow-up to examine the effects of the video-feedback training in study 3, were used to control for possible effects of that training in the present study.

The consultancy program consisted of three consultations in total, with two onsite consultations and a third consultation given per telephone. During the consultancy program, center directors used a self-assessment tool (the NCKO Quality Monitor; Gevers Deynoot-Schaub et al., 2009) to self-evaluate the quality of the child care environment in care groups. Based on the self-assessment and the ITERS-R/ECERS-R pretest scores, the consultant and center director together drew up an action plan for improvement which was tailored to the individual needs of the group: the consultancy specifically targeted items that were identified as 'weak areas' by the consultant and center director.

Results demonstrated no significant effects for the ITERS-R and ECERS-R total scores. The results for the total scores were nonetheless meaningful in terms of percentages of groups switching from inadequate to adequate care categories as defined for the ITERS-R and ECERS-R: 25% of the experimental groups improved from inadequate care to moderate care between pretest and follow-up on the ITERS-R and 23% of the groups on the ECERS-R. Furthermore, we found a significant improvement for the specific ITERS-R and ECERS-R items that had been targeted during the consultancy.

These promising, although moderate, outcomes demonstrate that the quality of the child care environment can be improved using the consultancy program for center directors, and that this effect remains over time. The results also show that the effects of the program do not generalize to quality topics beyond those addressed during the consultancy, so that care should be taken to include all topics that deserve improvement in the intervention aims.

## 6.2 Conclusions

The results of this thesis can be summarized in the following main conclusions:

- Reliability and validity of the CIP scales in the Dutch child care context is supported, which means that the scales can be applied for assessment of child care quality in The Netherlands to get a detailed picture of caregivers' interactive skills, and provide important information on child care quality in addition to the ITERS-R and ECERS-R that measure more global quality of the child care environment.

- Caregivers in The Netherlands score, on average, relatively high on the more basic CIP skills (i.e., sensitive responsiveness, respect for autonomy, and structuring and limit setting), but the range of scores is considerable and also includes inadequate scores. Caregivers' scores on the more educational CIP skills (i.e., verbal communication, developmental stimulation, and fostering positive peer interactions) are clearly lower and mostly in the inadequate range.
- It is important to measure and train caregiver interactive skills at an individual level, because caregivers have a personal and individually determined CIP interactive skills profile.
- The CIP training effectively strengthens caregivers' interactive skills. Additional booster sessions may be needed to increase the effectiveness of the training, by further improving the caregivers' educational skills.
- Global process quality in The Netherlands as assessed with the ITERS-R and ECERS-R is at a low to medium level and needs improvement. Therefore, it is important to continue monitoring child care quality.
- The consultancy program for center directors to improve global child care quality is effective, but only for items that are specifically targeted during the consultancy. Effects of the program do not generalize to topics beyond those addressed during the consultancy.

### 6.3 General discussion

Several issues regarding this thesis deserve further discussion. First, it is an important implication of the studies examining the effects of the intervention program that the current set-up of the intervention program with both the CIP training and the consultancy program for care groups should be retained, because the two components of the intervention program evidently affect different aspects of child care quality. The complete intervention program has a double focus, with one part aiming at improving the quality of the caregiver-child interaction and the other part aiming at improving the broader quality of child care environment (i.e., space, furnishings, materials, and program structure). Effects of the two intervention components were measured with the CIP scales and the ITERS-R/ECERS-R. Although both aspects of child care quality are related – validation of the CIP scales (see study 1) yielded a positive and significant correlation between the CIP scales and the ITERS-R/ECERS-R ( $r = .38$  for the CIP total score and the ITERS-R/ECERS-R total score) – the experimental studies for the two parallel parts of the intervention program nevertheless demonstrated that correlations between the gain scores obtained for the CIP training and the ITERS-R/ECERS-R gain scores obtained for the consultancy program were low and non-significant (not mentioned in study 3 and 4). Evaluation of both parts of the intervention program thus demonstrated that improvement of caregiver interactive skills did not necessarily lead to improvement of the global

quality of the child care environment, or vice versa. This is a clear indication that it is important to retain the double focus approach of the intervention program. Merely choosing one of the two intervention components is only justified when either one of the quality measures (caregiver interactive skills or the global quality of the child care environment) is notably lower than the other quality measure.

A second point that deserves attention is the finding that the effects of the CIP training were more apparent than the effects of the consultancy program. Obviously, the two complementary programs have a different focus and method. The most evident strength of the CIP training is the video feedback method. Previous studies have already endorsed the use of video feedback for training programs targeting parent or caregiver behavior (Bakermans-Kranenburg, Van IJzendoorn, & Juffer, 2003; Fukkink & Lont, 2007; Fukkink & Tavecchio, 2010; Groeneveld, Vermeer, Van IJzendoorn, & Linting, 2011). Given that the consultancy program does not focus on the center director's interactive behavior, but rather on bringing about changes in environmental characteristics such as space, furnishings, and materials, video feedback did not directly apply for the consultancy program. Nevertheless, it would be interesting to incorporate some visual tools in the consultancy. For instance, photographs of good practice examples for space, furniture, and materials could be very useful to visualize the intended purpose of the program.

Another important difference between the two parts of the intervention program is that the intensity of the consultancy program was lower than the intensity of the CIP training. The CIP training consisted of 5 sessions, while the consultancy program merely consisted of two onsite consultations and a third more 'distant' session by telephone. Although prior research is not clear about the effects of program intensity in the case of consultancy programs (Zaslow et al., 2010), the lower intensity of the consultancy program may have contributed to the less favorable outcomes as compared to the CIP training.

A final difference between the two intervention program parts that may explain their difference in effectiveness is that the trainees in the CIP training (i.e., the caregivers) could directly realize their intervention aims in their own behavior in interaction with the children. In the consultancy program, however, center directors had to attain their aims in a more indirect way. They were asked to improve the child care environment in the care group of the caregivers, whose cooperation was required to realize many of the intervention targets. Hence, instructing the center directors on how to improve environment characteristics was only a first step in the actual improvement process; the second step was for center directors to realize these changes in close collaboration with the caregivers in the care group. For example, center directors could purchase new materials, but implementing these materials eventually boiled down to what the caregivers in the care group did with them. Obtaining adequate to good scores for materials, according to the ITERS-R/ECERS-R, requires not only availability of the materials but also involvement of the caregiver when children play with the materials. The main reason that the consultancy was nevertheless directed at center directors and not caregivers was be-

cause caregivers' influence on the broader quality of the child care environment is limited, given that it is the center director who decides about provisions, materials, and the program. For future research, however, we would suggest to also involve caregivers in the consultancy process, provided that this does not lay a heavy burden on the caregivers. To prevent an overload for caregivers, the two training programs could be run sequentially instead of simultaneously. The order in which the programs are offered does not seem very important. An advantage of starting the intervention with the CIP training might be that there is room for the suggested additional booster sessions of the CIP training (see study 3, chapter 4) during the subsequent consultancy program. Altogether, the abovementioned differences between the CIP training and the consultancy program may explain why the effects of the CIP training appear more favorable than the effects of the consultancy program.

A third point of discussion concerns initial vocational education for caregivers. Caregiver education in The Netherlands evidently does not pay enough attention to the educational CIP skills. Study 1 and 2 (Chapter 2 and 3) of this thesis demonstrated that although caregivers in The Netherlands score, on average, relatively high on the more basic CIP skills (i.e., sensitive responsiveness, respect for autonomy, and structuring and limit setting), scores on the more educational CIP skills (i.e., verbal communication, developmental stimulation, and fostering positive peer interactions) are mostly in the inadequate range. Thus, from the perspective of the child, caregiver education falls short; students are not well enough instructed on how to adequately stimulate children's language, cognitive, creative, motor, and social development. In recent years some adaptations in Dutch caregiver education have been made in response to the acknowledged shortcomings of the previous vocational training. As from September 2014, caregiver education in The Netherlands will get a new profile with a specific focus on all six CIP skills (Calibris, 2014). Of course, this requires specific expertise and new skills of teacher educators to implement this profile and therefore a straightforward next step would be to implement the CIP training in caregiver education. The NCKO is currently developing a *train the trainer* program of the CIP training for teacher educators in caregiver education. Effects of the implementation of this adapted CIP training (including the train the trainer) in caregiver education will be examined in future research.

A final point of discussion pertains to the international relevance of the studies described in this thesis, considering that they are based on the NCKO Quality Model (see Chapter 1 of this thesis), which may reflect Dutch values and characteristics of the Dutch child care context. Although there are evidently culturally specific values for defining, measuring, and improving child care quality, recent interest in the NCKO research project from Norway suggests that there is broader support for the Dutch view on child care quality and the corresponding line of research. The complete NCKO research plan has been adopted by Norwegian colleagues and their proposal (which also includes some possible adaptations to the Norwegian context) has been granted by the Norwegian government for a period 2012-2020 (Bjørnes-tad, Gulbrandsen, Johansson, & Os, 2013; Johansson, 2012). As a result, the CIP

scales have now been translated into Norwegian and the first promising observation results suggest that the CIP scales can be satisfactorily applied in the Norwegian child care context.

Interest for the CIP scales has also been shown by researchers in the United States, and plans are being made for a collaborative study to examine how the CIP scales are related to the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008; Roorbach Jamison, Cabell, LoCasale-Crouch, Hamre, & Pianta, 2014) – which assesses teacher-child interactions on several dimensions, including some that are similar to the CIP scales. Both examples of international collaboration demonstrate that there is international interest in the NCKO research and future research should explore whether the NCKO Quality Model is also applicable in other countries with different child care contexts.

#### **6.4 Limitations and future directions**

Beyond the limitations that are already mentioned above and those mentioned in the chapters reporting on the separate studies, two limitations and future directions deserve further attention. First, the NCKO Quality Monitor was nationally distributed with government funding at the level of child care organizations in 2009. Thus, it is possible that center directors with groups allocated to the control condition had already worked with the monitor before the intervention program had started off, which may have acted as a confounder in the effect studies. To check for this, center directors in the experimental and control condition were asked whether they had heard about the Quality Monitor, and, if so, whether they had been working with (parts of) the Quality Monitor. About 80 percent of the center directors indicated that they had heard about the monitor. Eight center directors answered that their organization had received the monitor, yet only two directors indicated that they had actually used the monitor: one center director with groups that were allocated to the control condition, and one center director with groups that were allocated to the experimental condition. We compared pretest scores of the groups of which the center directors had indicated that they had worked with the monitor (9 groups with 14 caregivers), with the groups that had not previously worked with the monitor (59 groups with 124 caregivers). Results indicate that there were no significant differences in ITERS-R and ECERS-R or CIP pretest scores between the groups for which the center directors had or had not worked with the monitor previously.

Second, the intervention was carried out in mixed-age groups only. Although mixed-age groups cover the child care age range from 0- to 4-year-olds, it cannot be automatically assumed that the effects of the intervention program can be generalized to infant and preschool groups. Therefore, future research in samples with homogeneous infant and preschool groups in child care centers is recommended to draw definite conclusions about the effects of the intervention program in other

group types. This issue is also interesting from an international perspective, given that mixed-age groups are typical for the Dutch situation but are less common in other countries.

## References

- Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin*, *129*, 195-215. doi: 10.1037/0033-2909-129.2.195
- Bjørnstad, E., Gulbrandsen, L., Johansson, J. E., & Os, E. (2013). *Metodiske idealkrav og nødvendige tilpasninger: Foreløpig tilstandrapport fra prosjektet* [Better provisions for Norway's children in ECEC: A study of children's well-being and development in ECEC]. Oslo, Norway: Høgskolen I Oslo og Akershus (HiOA).
- Calibris (2014). *Kwaliteitsdossier MBO, Pedagogisch Werk*. [Quality file Intermediate Vocational Training, Pedagogical Work]. Bunnik, The Netherlands, Calibris. Retrieved from [http://www.calibris.nl/-Onderwijsdocumenten/Kwalificatiedossiers/KD-s-2014-2015-\(concept\)/Kwalificatiedossier-Pedagogisch-Werk](http://www.calibris.nl/-Onderwijsdocumenten/Kwalificatiedossiers/KD-s-2014-2015-(concept)/Kwalificatiedossier-Pedagogisch-Werk)
- Fukkink, R. G., & Lont, A. (2007). Does training matter? A meta-analysis and review of caregiver training studies. *Early Childhood Research Quarterly*, *22*, 294-311. doi: 10.1016/j.ecri.2007.04.005
- Fukkink, R. G., & Tavecchio, L. W. C. (2010). Effects of video interaction guidance on early childhood teachers. *Teaching and Teacher Education*, *26*, 1652-1659. doi: 10.1016/j.tate.2010.06.016
- Gevers Deynoot-Schaub, M., Fukkink, R., Riksen-Walraven, M., De Kruijf, R., Helmerhorst, K., & Tavecchio, L. (2009). *De NCKO-Kwaliteitsmonitor* [The NCKO Quality Monitor]. Amsterdam, The Netherlands: SWP.
- Groeneveld, M. G., Vermeer, H. J., Van IJzendoorn, M. H., & Linting, M. (2011). Enhancing home-based child care quality through video-feedback intervention. *Journal of Family Psychology*, *25*, 86-96. doi: 10.1037/a0022451
- 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.
- Johansson, J. (2012). *Better provision for Norway's children in ECEC. A study of children's well-being and development in ECEC, and new tools for quality evaluation*. Application to The Research Council of Norway. Oslo, Norway: Oslo and Akershus University College for Applied Sciences.
- Oberhuemer, P., Schreyer, I., & Neuman, M. J. (2010). *Professionals in early childhood education and care systems. European profiles and perspectives*. Opladen & Farmington Hills: Barbara Budrich Publishers.
- Palsha, S. A., & Wesley, P. W. (1998). Improving quality in early childhood environments through on-site consultation. *Topics in Early Childhood Special Education*, *18*, 243-253.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System [CLASS] Manual: Pre-K*. Baltimore, MD: Brookes Publishing.
- Roorbach Jamison, K., Cabell, S. Q., LoCasale-Crouch, J., Hamre, B. K., & Pianta, R. (2014). CLASS-Infant: An observational measure for assessing teacher-infant interactions in center based care. *Early Education and Development*, *25*, 553-572. doi: 10.1080/10409289.2013.822239
- Zaslow, M., Halle, T., & Tout, K. (2011). Differing purposes for measuring quality in early childhood settings: aligning purpose with procedures. In: M. Zaslow, I. Martinez-Beck, K. Tout, & T. Halle (Ed.), *Quality Measurement in Early Childhood Settings* (p. 389-410). Baltimore, MD: Paul H. Brookes Publishing Co.