



## Improving quality of the child care environment through a consultancy programme for centre directors

Katrien O. W. Helmerhorst, Ruben G. Fukkink, J. Marianne A. Riksen-Walraven, Mirjam J. J. M. Gevers Deynoot-Schaub & Louis W. C. Tavecchio

To cite this article: Katrien O. W. Helmerhorst, Ruben G. Fukkink, J. Marianne A. Riksen-Walraven, Mirjam J. J. M. Gevers Deynoot-Schaub & Louis W. C. Tavecchio (2017) Improving quality of the child care environment through a consultancy programme for centre directors, *International Journal of Early Years Education*, 25:4, 361-378, DOI: [10.1080/09669760.2017.1321528](https://doi.org/10.1080/09669760.2017.1321528)

To link to this article: <https://doi.org/10.1080/09669760.2017.1321528>



© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 08 May 2017.



Submit your article to this journal [↗](#)



Article views: 683



View Crossmark data [↗](#)

## Improving quality of the child care environment through a consultancy programme for centre directors

Katrien O. W. Helmerhorst<sup>a</sup>, Ruben G. Fukkink<sup>a</sup>, J. Marianne A. Riksen-Walraven<sup>b</sup>,  
Mirjam J. J. M. Gevers Deynoot-Schaub<sup>c</sup> and Louis W. C. Tavecchio<sup>a</sup>

<sup>a</sup>Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, Netherlands; <sup>b</sup>Department of Developmental Psychology, Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, Netherlands; <sup>c</sup>Kohnstamm Institute, Amsterdam, Netherlands

### ABSTRACT

This study examined the effects of a newly developed on-site consultancy programme to improve global quality of the child care environment in non-parental child care centres for 0- to 4-year-old children as measured with the ITERS-R/ECERS-R. Using a randomised controlled trial with a pretest, posttest, and follow-up test, we compared 35 experimental group with 33 control group. The consultancy programme comprised three consultations in total. Analysis on the items that were specifically targeted during the consultancy showed a significant improvement on these targeted items between pretest and posttest and between posttest and follow-up. The effect of the consultancy programme on the total scores (including the non-targeted items) was not significant.



### ARTICLE HISTORY

Received 2 September 2016  
Accepted 5 March 2017

### KEYWORDS

Centre-based child care; consultancy; randomised controlled trial; global quality of the child care environment; ITERS-R; ECERS-R

Ample research has shown that the quality of early non-parental child care contributes to children's socioemotional and cognitive development (see Belsky et al. 2007; Vandell and Wolfe 2000; Vandell et al. 2010 for an overview). Unfortunately, repeated quality assessments in 1995, 2001, 2005, and 2008 in nationally representative samples of child care centres indicate that the quality of child care for 0- to 4-year-olds in Dutch child care centres has been steadily decreasing over the last decades (De Kruif et al. 2009; Helmerhorst et al. 2015; Vermeer et al. 2008). In these studies, the quality of the child care environment (QCCE) was measured using the Infant/Toddler Environment Rating Scale (ITERS-R; Harms, Cryer, and Clifford 2003) and the Early Childhood and Environment Rating Scale (ECERS-R; Harms, Clifford, and Cryer 1998). The ITERS-R/ECERS-R total score decreased from 4.8 on a 7-point scale in 1995 to 4.3 in 2001, 3.6 in 2005, and 3.0 in 2008. These scores are low both in absolute terms and from an international perspective (De Kruif et al. 2009; Helmerhorst et al. 2015; Vermeer et al. 2008). In 2008, none of the groups scored in the category good to excellent, 49% scored moderate, and more than half (51%) of the groups scored inadequate (see Helmerhorst et al. 2015 for a discussion of the results). Previous research has demonstrated that quality of care as measured with the

**CONTACT** Katrien O. W. Helmerhorst  K.O.W.Helmerhorst@uva.nl  Research Institute of Child Development and Education, University of Amsterdam, PO Box 94208, 1090 GE Amsterdam, Netherlands

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

ITERS-R/ECERS-R is related to children's development (e.g. Vandell and Wolfe 2000). Altogether, this clearly underscores the need for improving quality of care in Dutch child care centres for 0- to 4-year-olds. The present paper describes the effects of a consultancy programme for directors of child care centres aimed to improve the QCCE as measured by the ITERS-R and ECERS-R.

The consultancy programme described in the present paper was part of a more comprehensive intervention programme to improve the quality of care in Dutch child care centres. We defined high-quality care as care that contributes to children's well-being and development (see e.g. Layzer and Goodson 2006). Children develop through their direct interactions with caregivers, other children in the group, and the available materials. These direct interactions are defined as process quality (e.g. Layzer and Goodson 2006; Vandell and Wolfe 2000), which is generally acknowledged as the core of the quality provided in child care groups. The complete intervention had two main goals: (1) to improve the quality of caregiver-child interactions and (2) to improve the more global quality characteristics of the child care environment, i.e. the physical environment such as space, furnishings, materials, and program structure. Henceforth, QCCE as measured with the ITERS-R/ECERS-R will be referred to as 'global quality' in this paper.

An important question in developing the intervention programme was at which person (s) in the child care centre the intervention should be directed. In the Netherlands, centre directors are the managers of the child care centre and responsible for finances, planning, and pedagogical policy and quality by supervising the caregivers of the care groups. Based on this, we believed that centre directors would be the key stakeholders to improve the more global quality characteristics of the child care environment by taking part in a consultancy programme. Improving the quality of the caregiver-child interactions, on the other hand, was thought to be most effective by means of a skills training directly provided to the caregivers of the children. Therefore, we developed the 5-week Caregiver Interaction Profile (CIP) training for caregivers in addition to the consultancy programme for the centre directors. The CIP training and its effects are described in a separate paper (Helmerhorst et al. 2016). Taken together, the complete intervention programme consisted of two components that ran simultaneously: (1) a consultancy programme directed at the centre directors to improve the more global quality characteristics of the child care environment, i.e. the physical environment such as space, furnishings, materials, and in addition everyday program structure, and (2) a video-feedback training directed at caregivers to strengthen their skills in interacting with the children (CIP training).

In the present study, we examined the effects of the first component of the intervention programme, namely the on-site QCCE consultancy programme directed at centre directors to improve global QCCE as measured with the ITERS-R/ECERS-R.

## **Results of previous studies on effectiveness of consultancy in child care settings**

Because our intervention was directed at centre directors, we started by searching for prior interventions aimed at improving global quality and directed at centre directors. Up until now, earlier interventions focusing on improving global child care quality through on-site consultancy (all conducted in the United States) have come up with mixed results. We found only one study, by Bloom and Sheerer (1992), which evaluated a 16-month

programme that focused on a broad set of domains (i.e. personal and professional self-knowledge, child development, organisational theory, leadership, parent relations), with quality of the classroom being one of the outcome variables. Overall classroom quality was significantly higher in target groups which received the programme than in comparison groups without intervention.

We found several other intervention studies that were directed at caregivers of a care group. In a review study, Zaslow et al. (2010) discuss five studies that are specifically focused on strengthening global QCCE. First, a study by Fiene (2002) describes a 4-month intensive mentoring programme for infant caregivers delivered by an early childhood professional. Unfortunately, little information is given about the content of the mentoring programme. No effects were found on global quality of the classroom environment as measured with the ITERS. In a second study, Kontos, Howes, and Galinsky (1996) report the effects of an on-site training for family child care providers (including workshops and home visits) that modestly improved global quality in two out of three sites as measured with the FDCRS (Harms and Clifford 1989). Modest effects were also found in a third study by Campbell and Milbourne (2005). In this study, infant-toddler caregivers received both a group training course (five 3-hour sessions) on issues related to children with special needs, infant and toddler development, learning, and socialisation, and three 1-hour consultation visits. Caregivers used a self-assessment instrument to indicate targets for improvement that were related to the ITERS subscales space and furnishings, personal-care routines, activities, interactions, program structure, and adult needs. Results showed no significant improvement on the ITERS subscales. Finally, two studies describing on-site interventions demonstrated significant improvement on global quality in centre-based infant-toddler and preschool classrooms (Palsha and Wesley 1998; Wesley 1994). Sample sizes in these studies were small, and the design did not include a control group. Both studies used the ITERS and ECERS rating scales as the basis for the intervention; the intervener trained the child care provider to use the rating scales themselves. The intervener and child care provider together drew up an action plan for improvement. The number of visits ranged between 10 and 14 over a period of 10–12 months. The study by Palsha and Wesley (1998) found a significant improvement by more than a half scale point on the ITERS and ECERS total score.

Finally, the Quality Interventions for Early Care and Education (QUINCE) study was the only study we found with a randomised controlled trial (RCT) design (Bryant et al. 2009). An RCT is a controlled experimental design with random assignment of participants to the different experimental conditions (intervention versus control group), which allows to determine effects of the intervention while controlling for other variables. This design is generally considered as the golden standard in intervention research (e.g. Schulz, Altman, and Moher 2010).

The QUINCE study evaluated effectiveness of the Partnership for Children (PFI) individualised on-site consultation programme with an average of 19.3 visits in total. This consultation model (see below under 'Development of the Present Consultation Model' for a detailed description of the PFI programme) was also used in the aforementioned studies of Palsha and Wesley (1998) and Wesley (1994). The QUINCE study compared PFI to typical services for quality enhancement programmes which the control group received, with an average of 6.7 visits. Results from the QUINCE study indicated no differences between the experimental group that received PFI consultation and the control group

in centre-based care with a regular quality improvement policy: both groups improved over time on the ECERS-R factors teaching and interactions and provisions for learning.

Taken together, results from previous U.S. studies on effectiveness of on-site interventions aimed at improving global child care quality are promising but far from conclusive. From the abovementioned studies examining intervention effects, merely one used a randomised controlled design. To draw more definite conclusions about the effects of this type of intervention, controlled experimental research is needed.

A relevant question is whether the gains of consultancy programmes have been adequately captured in previous research. QCCE comprises many facets, which are aggregated in global quality measures, whereas the strength of many consultancy programmes is that they are tailored to the needs, aims, and preferences of the child care provider and are thus highly flexible in the choice of specific targets for improvement (Wesley and Buysse 2004). The aforementioned studies have used global assessment instruments with a wide scope to measure training outcomes and may, therefore, not capture the specifically targeted elements of the intervention (Zaslow et al. 2010; Zaslow, Halle, and Tout 2011). Thus, there may be a discrepancy between the highly individualised intervention programmes and the broad assessment instruments that have been used to evaluate the effectiveness of this type of consultancy interventions. This emphasises the importance of taking into account the specific purpose of measurement in choosing an assessment instrument (see also Snow and Van Hemel 2008; Zaslow, Halle, and Tout 2011), because when measurement instruments do not closely align with the content of the intervention, intervention effects do not become visible (Fukkink and Lont 2007). Therefore, we evaluated the effectiveness of the present consultancy programme on the items that were specifically targeted during the consultancy.

## Development of the present consultation model

The PFI on-site consultation model that was used in the earlier mentioned QUINCE study (Bryant et al. 2009) and in the studies by Palsha and Wesley (1998) and Wesley (1994) showed positive results. The model describes a framework with several phases in which the consultant and consultee work in productive collaboration on overall quality improvement with a strong focus on tailoring the consultancy to the individual needs of the consultee or specific centre. We decided to work with elements of the PFI model, which served as a starting point for the set-up of our QCCE consultancy programme. As can be seen in Figure 1 (right side), the PFI model includes eight steps in total (Bryant et al. 2009; Palsha and Wesley 1998). The first step in the consultation process according to this model is *establishing the consulting relationship*. The consultant starts with gaining the trust of the consultee by creating a collaborative basis and together they develop an action plan for improvement. The second step is to *provide training on the scale* to the consultee. During the third step, the consultant and consultee *jointly assess needs*; the consultant focuses on determining the factors needed for improvement (for example more play and learning material that are accessible to the children). Both the consultant and the consultee administer a broad quality measure as the ECERS-R/ITERS-R scales. Self-assessment by the consultee is important because it shows staff that their input in the consultation process is vital; staff is actively involved in setting goals and standards for their future professional development; and it allows staff to get insight in the current quality level of their childcare centre (Wesley

<u>Present QCCE consultation model</u>	<u>PFI consultation model</u>
<b>Pretest</b> Independent observer scores rating scale	<b>Step 3</b> a. Consultant administers rating scale
<b>Session 1</b> <ul style="list-style-type: none"> <li>Establishing the consulting relationship</li> <li>Explanation on how to administer self-assessment</li> </ul>	<b>Step 1</b> Establishing the consulting relationship  <b>Step 2</b> Provide training on the scale
<b>Between session 1 &amp; session 2</b> Consultee administers self-assessment	<b>Step 3</b> Jointly assess needs: b. Consultee administers rating scale c. Consensus meeting to assess needs
<b>Session 2</b> <ul style="list-style-type: none"> <li>Develop action plan</li> <li>Consultant provides advice for implementation action plan</li> </ul>	<b>Step 4</b> Develop action plan  <b>Step 5</b> Implementation of the action plan
<b>Session 3</b> Consultant follows up on implementation action plan and identifies pending questions	<b>Step 8</b> Identify future needs
<b>Posttest</b> <ul style="list-style-type: none"> <li>Independent observer scores rating scale</li> <li>Consultee fills in evaluation form</li> </ul>	<b>Step 6</b> Evaluate changes  <b>Step 7</b> Write a final report and hold summary conference
<b>Follow-up</b> Independent observer scores rating scale	

**Figure 1.** QCCE consultation model versus PYI consultation model.

1994). Altogether, this should provide better chances for maintaining any quality improvement over time when consultancy is completed. Subsequently, consultant and consultee collaboratively *develop an action plan* (for example, lowering the existing play and learning materials to a place where they are within view and reach for the children and to purchase additional play and learning materials). This phase highlights the often unique and group-specific nature of the intended quality improvement. Consultant and the consultee together work on strategies for improvement. The consultee plays an active role in this phase (for example, providing suggestions where and how to display the play and learning materials).

During the fifth step, *implementation of the action plan*, the consultant gives advice on how to implement the written plan of action. *Evaluate changes* is the sixth step of the model. The consultant administers the rating scale again to assess whether the desired and actual outcomes identified correspond with the original action plan (for example, the ITERS/ECERS subscale Activities is administered for the second time). The goal of the seventh step is to *write a final report and hold a summary conference* together. The consultee evaluates the assistance of the consultant. Finally, during the last step, *identify future needs*, additional needs are identified by the consultee.

Campbell and Milbourne (2005) suggest that a programme with a group training course and three 1-hour consultation visits may already be effective in enhancing global quality as measured with the ITERS/ECERS. Therefore, we decided to compress the PFI consultation model to a three-session consultancy programme. We believed that this was possible because of the more narrow focus of our QCCE consultancy programme; the present programme did not address caregiver–child interactions, whereas the interventions that used the PFI consultancy model did. Figure 1 presents an overview of the PFI consultation model (right side) and the present QCCE consultancy programme (left side) and shows how the three sessions of the QCCE programme are related to the eight steps of the PFI model. As can be seen in Figure 1, all steps of the PFI model are included in the three sessions of the QCCE consultancy, although in a slightly different order.

## Present study

The main aim of this study was to examine the effects of the newly developed on-site QCCE consultancy programme to improve the global QCCE for 0- to 4-year-old children as measured with the ITERS-R and ECERS-R. The consultancy programme focused on four quality domains, represented by the ITERS-R/ECERS-R subscales *space and furnishings*, *language*, *activities*, and *program structure*. Consultancy specifically targeted items that consultant and consultee together identified as ‘weak areas’ for a given care group of 0- to 4-year-olds. As mentioned in this Introduction, effects of consultancy programmes are generally assessed with global quality measures and thus may not capture the specifically targeted elements of the consultancy intervention. Therefore, we first tested the consultation effects focusing on the exact objectives by analysing the specific ITERS-R/ECERS-R items targeted during consultation. We expected a positive effect of the present QCCE programme on the items that were specifically targeted during the consultancy. Second, we also examined the effects of the QCCE consultation on the total score of the relevant ITERS-R/ECERS-R subscales, which also included the items that were not targeted during the consultancy – and which were therefore expected to be not (or less) affected by the consultation. Finally, we examined the effects of the QCCE consultancy on the items that were *not* targeted during the intervention.

## Method

### Participants and randomisation

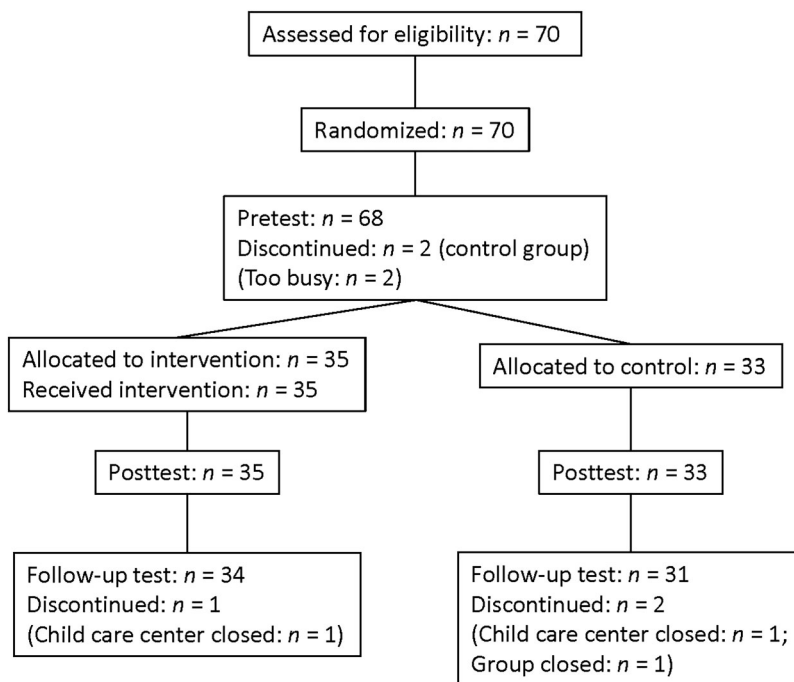
Child care groups in this study were recruited from child care centres in and around Amsterdam, the Netherlands. Centre directors responded to appeals in (digital)

newsletters and announcements on child care websites. Centres had to participate with an even number of groups to assign to the study. Furthermore, participating groups had to be mixed-age groups (0- to 4-year-olds), because we wanted to focus on groups with children across the whole possible age range and because mixed-age groups are more prevalent in the Netherlands than infant groups (0- to 2-year-olds) and preschool groups (2- to 4-year-olds) (see De Kruif et al. 2009; Helmerhorst et al. 2015).

Originally, the sample consisted of 70 child care groups: half of the groups were randomly assigned to the experimental condition (QCCE consultancy programme) and half were assigned to the control condition (no consultancy at all). Two groups assigned to the control condition dropped out after randomisation and before the pretest. The final sample included a total of 68 mixed-age groups from 33 child care centres with 35 groups in the intervention condition and 33 groups in the control condition. Across the 35 experimental groups, a total of 14 centre directors participated in the consultancy programme with 1 group, 6 centre directors with 2 groups, and 3 centre directors with 3 groups.

Figure 2 shows the flow chart of the participating groups per phase. Three groups (one in the experimental condition and two in the control condition) discontinued between posttest and follow-up because the child care centre or the group had been closed. None of the participating centre directors dropped out in the course of the intervention programme.

On average, child care centres were in existence for 12.4 years ( $SD = 9.45$ , range 0.5–30) and had 3.7 groups per centre ( $SD = 1.84$ , range 1–8). Centre directors were all female; on average, they were 42.8 years old ( $SD = 9.77$ , range 27–61), worked 28.3 hours a week ( $SD = 7.7$ ), and had 13.4 years ( $SD = 8.5$ ) of working experience in child care. About half (52%)



**Figure 2.** Flow chart of the study's progress in terms of groups.



of the centre directors had a bachelor degree, 28% completed regular vocational training, and 10% had a master degree. The remaining 10% only completed secondary school.

Multivariate analyses showed no significant differences at the pretest between the experimental and the control group for either the ITERS-R or ECERS-R subscales, and both groups were equivalent at the start of the study (Wilks'  $\lambda = .947$ ,  $F(4, 63) = 0.87$ ,  $p = .49$ , and Wilks'  $\lambda = .955$ ,  $F(4, 63) = 0.74$ ,  $p = .57$ ,  $\eta^2_{\text{partial}} = .045$ , respectively).

### **Design and general procedure**

We used a controlled study design with random assignment to the experimental and control condition. The dependent variables, ITERS-R and ECERS-R scores, were measured at pretest, at posttest directly after the intervention, and at a 3-month follow-up.

As noted in the Introduction, this study was part of a more comprehensive intervention programme; parallel to the present consultancy programme for centre directors aimed at improving global quality, we conducted the CIP training for professional caregivers that was specifically aimed at improving the quality of caregiver–child interactions as measured with the CIP scales (De Kruif et al. 2007; Helmerhorst et al. 2014). In the present study, we focused on the effects of the consultancy aimed at global QCCE as measured with the ITERS-R and ECERS-R and controlled for possible effects of the CIP intervention (see below under Measures: CIP scales).

An outline of the general procedures can be found in Figure 1 (left-hand column in the present QCCE consultation model). All groups in experimental and control condition were visited by a trained observer for pretest, posttest, and follow-up test. The visit lasted from about 8 am until about 3 pm. In order to obtain independent ratings, observers visited a care group only once and were blind to the experimental condition (QCCE consultancy programme or no consultancy). At the pretest, posttest, and follow-up, the observer scored global process quality for the group with the ITERS-R and ECERS-R, and interviewed one of the caregivers to collect additional information (see Measures) that was not available through observation at the end of the observation. In addition, the observer made video recordings of each caregiver, which were rated afterwards for a separate study examining the effects of the parallel video-feedback training to improve caregivers' interactive skills (see Helmerhorst et al. 2016). Two weeks after the pretest, the centre director was visited by the consultant for the first time. Two weeks after the last consultation, or 6 weeks after the pretest for the control groups, each group was visited for the posttest. Three months after the posttest, each group was visited for the follow-up observation. The posttest and follow-up visits were planned on the same day of the week as the pretest for an optimal comparison. At the pretest, centre directors also completed a questionnaire to collect individual background information (i.e. education and work experience).

### **Intervention programme**

#### **Experimental group**

The present consultancy programme was conducted by two consultants: one consultant had a master degree in developmental psychology and the other had a PhD in psychology. Both consultants had extensive experience in early years care and education and working

with the QCCE as evaluated with the NCKO Quality Monitor (Gevers Deynoot-Schaub et al. 2009; see below). Both consultants used a standardised manual containing a detailed step-by-step protocol for the content of each of the three sessions of the consultancy programme.

The consultancy programme comprised three consultations in total, with two on-site consultations, each lasting about 2 hours, and a third consultation given per telephone. During the first consultation, 2 weeks after the pretest, the consultant explained the general consultancy procedure and provided the centre director with background information of the research. As a first step in the consultancy process, the centre director and the consultant jointly chose targets for the consultancy programme by means of a self-assessment tool. For this self-assessment we used the NCKO Quality Monitor, which is a tool developed for professionals in child care centres to broadly self-evaluate the QCCE of the care groups with a checklist. The QCCE checklist is rated with 26 quality indicators on a dichotomous scale with a positive and a negative anchor and is based on the ITERS-R and ECERS-R items from the subscales space and furnishings, activities, language, and program structure. The positive anchor represents items that should be present (i.e. represented by a 'thumbs up' symbol) in the care group and the negative anchor represents items that should not be present (i.e. represented by a 'thumbs down' symbol). By specifying the negative items and positive items, the QCCE checklist provides the centre director with direct insight into a centre's stronger and weaker points and which items need improvement (for examples, see [Appendix 1](#)). During the first consultation, the consultant explained the centre directors how to administer the QCCE checklist. After the first consultation, centre directors were asked to complete in the QCCE checklist by themselves and mail it back to the consultant. All centre directors were able to administer the QCCE checklist by themselves.

To prepare for the second session, the consultant analysed both the QCCE checklist filled in by the centre director and the ITERS-R and ECERS-R scores from the pretest, and drew up an action plan based on both forms. The consultancy manual stated that ITERS-R/ECERS-R items with a score 3 or lower would be marked with priority on the action plan. Consultants were asked to carefully list the items that were targeted during the QCCE consultancy for each group, so that we could examine the effect of the consultancy on the specific ITERS-R/ECERS-R items targeted during consultation. The number of action items was 13.4, on average, with a minimum of 8 and a maximum of 20 action points ( $SD = 2.91$ ). In all groups consultancy concerned activities for blocks and math/numbers ( $n = 35$ ), in most groups it concerned child-related display of visual materials for children (i.e. posters, collages, children's artwork, pictures), and music/movement ( $n = 34$ ) (see [Table A1](#) in the [Appendix](#) for an overview of the items that were targeted during the consultancy). In general, items of the subscale program structure were targeted least. During the second consultation, about 2 weeks after the first, the consultant brought back the QCCE checklist filled in by the centre director together with the action plan. After careful consideration, the centre director and the consultant jointly determined the definitive action points and developed a plan for quality improvement for the group in the given areas. This way the consultancy was tailored to the individual and unique needs of the group. The consultant advised the centre director on how to implement the improvement plan. After the second visit, the centre director worked on the action items list. The third consultation was a final phone call about 2 weeks later to check with the centre director

whether the actions items were feasible, and to answer questions when needed. After the last consult, we sent an evaluation form about the consultant and the QCCE consultancy in general to the centre director.

### **Control group**

Groups that were assigned to the control condition received no intervention and were only contacted to schedule the pretest, posttest, and follow-up test.

### **Measures**

#### **ITERS-R/ECERS-R**

ITERS-R (Harms, Cryer, and Clifford 2003) and its equivalent for preschool classrooms, the ECERS-R (Harms, Clifford, and Cryer 1998), are widely used to measure process quality in child care groups. The ITERS-R was developed for use in groups in which more than 50% of the children is under the age of 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 both ITERS-R and ECERS-R in each group at the pre-, posttest, and follow-up to capture quality for both infants and toddlers in the participating mixed-aged groups (0- to 4-year-olds). Both the ITERS-R and the ECERS-R include 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. Items are rated on a 7-point scale with descriptors for the scores 1 (inadequate), 3 (minimal), 5 (good), and 7 (excellent). We only used the subscales that were targeted in the consultancy programme, i.e. space and furnishings, language, activities, and program structure. A total score of the four subscales space and furnishings, language, activities, and program structure (20 items for the ITERS-R and 24 items for the ECERS-R) was computed by averaging item scores across the four subscales. Internal consistency for the ITERS-R (Cronbach's alpha) was .77 at the pretest, .63 at the posttest, and .77 at the follow-up test. For the ECERS-R, Cronbach's alpha was .70 at the pretest, .69 at the posttest, and .78 at the follow-up test.

Prior to data collection, eight observers were trained to use both the ITERS-R and the ECERS-R. Each observer visited at least 4 locations (range 4–6) supervised by an expert trainer, followed by an item-by-item debriefing. The average interobserver agreement (i.e. intraclass correlations) during training was .84, on average, ranging from .80 to .88. Observers were blind to the experimental condition of the group (consultancy programme or control group).

#### **CIP scales**

Quality of caregiver–child interactions was rated from the videotaped episodes with the CIP scales (De Kruif et al. 2007; Helmerhorst et al. 2014); the scores were used here to control for possible effects of the parallel caregiver interaction training (CIP training). The CIP scales reflect six caregiver interactive skills: *sensitive responsiveness*, *respect for autonomy*, *structuring and limit setting*, *verbal communication*, *developmental stimulation*, and *fostering positive peer interactions*. Each of the CIP scales is rated on a single 7-point scale (7 = very high, 6 = high, 5 = moderate/high, 4 = moderate, 3 = moderate/low, 2 = low, 1 = very low). For a more comprehensive description of the CIP scales, see De Kruif

et al. (2007) and Helmerhorst et al. (2014). Per care group a CIP-total score was calculated by first averaging the scores of the six skills per caregiver and then averaging the scores of all caregivers per group. Next, a gain score was computed by calculating the difference between the pretest and follow-up CIP-total scores. This gain score was included in the analyses to control for possible effects of the parallel video-feedback training for caregivers that aimed to improve their interactive skills as measured with the CIP scales (see Helmerhorst et al. 2016).

### **Participants' evaluation of the intervention**

Centre directors were asked to rate their satisfaction with the consultant and the consultancy programme directly after the intervention (i.e. at posttest) to obtain the programme's social significance. Centre directors completed the Consultant Evaluation Form (CEF; Erchul 1987), which consists of 12 items rated on a 7-point Likert scale (1 = *strongly disagree*, 7 = *strongly agree*). The CEF measures centre directors' perceptions of the consultant's effectiveness and their satisfaction with the consultant. We added four items to the CEF to rate overall satisfaction with the consultancy programme. Higher ratings reflected more favourable evaluation of the consultant's effectiveness and more satisfaction with the consultancy programme.

## **Results**

Table 1 shows the descriptives for the ITERS-R and ECERS-R pretest, posttest, and follow-up scores across the four relevant subscales and the total score for both the consultancy group and the control group.

### **Effects of the consultancy programme on targeted items**

First, we examined the effect of the consultancy specifically on the ITERS-R and ECERS-R items that had been targeted during consultancy. For that purpose, we computed an

**Table 1.** Mean subscale and total scores for the ITERS-R and ECERS-R in the experimental and control group at pretest, posttest, and follow-up.

Variable	Experimental group			Control group		
	Pretest (n = 35) M (SD)	Posttest (n = 35) M (SD)	Follow-up (n = 34) M (SD)	Pretest (n = 33) M (SD)	Posttest (n = 33) M (SD)	Follow-up (n = 31) M (SD)
<i>ITERS</i>						
Space and furnishings	2.84 (0.51)	3.25 (0.61)	3.30 (0.76)	3.01 (0.72)	3.12 (0.51)	3.16 (0.59)
Language	3.44 (0.80)	3.80 (1.07)	4.21 (1.07)	3.26 (0.93)	3.15 (0.95)	3.49 (1.22)
Activities	2.18 (0.51)	2.34 (0.67)	2.49 (0.70)	2.21 (0.65)	2.45 (0.49)	2.49 (0.62)
Program structure	3.99 (0.88)	4.25 (0.98)	4.58 (1.09)	4.09 (1.11)	3.62 (0.78)	4.32 (1.08)
Total 4 subscales <sup>a</sup>	2.81 (0.45)	3.08 (0.56)	3.26 (0.67)	2.86 (0.61)	2.90 (0.36)	3.08 (0.55)
<i>ECERS</i>						
Space and furnishings	3.09 (0.46)	3.34 (0.53)	3.59 (0.73)	3.26 (0.59)	3.23 (0.41)	3.52 (0.56)
Language	3.19 (0.69)	3.62 (0.92)	3.88 (0.97)	3.13 (0.79)	3.11 (0.75)	3.41 (0.85)
Activities	1.90 (0.36)	2.04 (0.50)	2.27 (0.57)	1.87 (0.36)	1.99 (0.33)	2.13 (0.50)
Program structure	3.45 (0.72)	3.71 (0.66)	3.74 (0.88)	3.52 (0.83)	3.13 (0.60)	3.67 (0.74)
Total 4 subscales <sup>a</sup>	2.70 (0.35)	2.95 (0.46)	3.16 (0.58)	2.75 (0.43)	2.73 (0.29)	3.00 (0.43)

Note: <sup>a</sup>Total 4 subscales = Total score for items of space and furnishings, language, activities, and program structure subscales.

aggregated variable for each experimental group for the ITERS-R and ECERS-R by averaging scores on only those items that had been targeted for the care group in question. We hypothesised that this variable, which is referred to as ‘targeted items’, was more sensitive to intervention effects, because the exact focus of the consultancy varied considerably between groups and was, therefore, related to different ITERS-R and ECERS-R items. This analysis only applied to the experimental group (there were no targeted items in the control group). We used multilevel repeated-measures analysis separately for the ITERS-R and ECERS-R. The dependent variable in both models was the aggregated targeted items variable, time was the repeated-measures factor with a pretest, posttest, and follow-up, and again the gain score of the parallel CIP training was included as a control variable.

Table 2 shows the results for the two multilevel models. We found a significant increase in the targeted items score for both the ITERS-R and ECERS-R. Results showed a significant increase between the pretest and the posttest and between posttest and follow-up for both the ITERS-R and ECERS-R items targeted during the consultancy ( $p < .001$  for all tests). As can be seen from the estimated means in Table 2, ITERS-R items targeted during the consultancy improved, on average, by 0.36 point between pretest and posttest and an additional 0.20 point between posttest and follow-up test. ECERS-R item scores improved, on average, by 0.30 point between pretest and posttest and an additional 0.30 point between posttest and follow-up test. The positive effects of the consultancy on the targeted items thus remained after three months and showed a small, but significant additional increase between posttest and follow-up test.

### **Effects of the consultancy programme on the total ITERS-R and ECERS-R scores**

We also examined the effects of the QCCE programme on the total ITERS-R and ECERS-R scores. We conducted multilevel analysis to take into account the hierarchical data structure with measurements at three time points (pretest, posttest, follow-up; Level 1) per group, and groups nested within centre directors (Level 2). Two separate models were analysed, one with the ITERS-R total score and one with the ECERS-R total score as the dependent variable, with ‘group’ as the between-subjects variable (a dummy variable: 0 = control group, 1 = experimental group) and time (pretest, posttest, follow-up) as the within-subjects variable. The gain score of the parallel intervention was included as a covariate in the model. We found no significant Group  $\times$  Time interaction effect for the

**Table 2.** Consultancy effects on ITERS-R and ECERS-R items targeted in experimental groups (multilevel analysis,  $n = 35$ ).

	ITERS-R items targeted		ECERS-R items targeted	
	Estimate	SE	Estimate	SE
<i>Fixed parameters</i>				
Intercept (pretest)	2.005**	0.103	2.075**	0.087
Posttest	0.364*	0.121	0.298**	0.104
Follow-up	0.557**	0.126	0.601**	0.117
Gain score parallel CIP-Training	-0.108	0.118	-0.084	0.095
Deviance	175.381		145.205	

Notes: \* $p < .05$ .

\*\* $p < .01$ .

consultancy programme, indicating that the effect of the consultancy on the ITERS-R and ECERS-R total scores was not significant.

Finally, we checked if there was a change over time for the items that were *not* targeted during the consultancy to get a more detailed picture of the outcomes of the QCCE consultancy. As expected, we found no significant increase for the items that were not targeted for either the ITERS-R or ECERS-R.

### **Participants' evaluation of the intervention**

Centre directors were on average very satisfied with the consultant ( $M = 5.98$ ,  $SD = 0.78$ , range = 3.92–7.00 on a 7-point scale). The items 'The consultant was generally helpful' and 'The consultant was a good listener' were rated highest by centre directors ( $M = 6.37$  and  $M = 6.42$ , respectively). Centre directors rated 'The consultant viewed her role as a collaborator rather than as an expert' lowest ( $M = 5.16$ ). In addition, centre directors were also very satisfied with the QCCE programme in general ( $M = 5.72$ ,  $SD = 0.72$ , range = 4.75–7.00) and rated the items 'The consultancy meetings were useful' and 'Pedagogical quality in the child care group has improved through the consultancy meetings' highest ( $M = 6.26$  and  $M = 5.67$ , respectively). On average, lowest scores were given to 'I became more competent through the consultancy meetings' ( $M = 5.32$ ).

### **Discussion**

The controlled experimental evaluation of our on-site consultancy programme for centre directors that aimed to improve quality of the environment in child care groups with a broad age range (0- to 4-year-olds) demonstrated a significant positive effect of the intervention for the items targeted during the consultancy. This suggests that QCCE improved during the intervention and this observed improvement remained at follow-up (three months after the posttest). It should be noted that the improvement was specific and was directly linked to the exact focus (targeted items) of the consultation, which varied significantly between the different groups. Hence, results of our consultancy intervention were not found in the analysis of the ITERS-R and ECERS-R total scores of the four subscales. However, a carefully constructed measure that matched the specific consultancy focus with the targeted items from these holistic measures was able to capture the subtle effects of the QCCE consultancy. In addition, analysis of the items that were not targeted showed there was no improvement over time for topics that were not addressed during the consultancy. This demonstrates that effects were indeed restricted to the specific topics that were addressed during the consultancy. Finally, centre directors were on average very satisfied with the consultant and the QCCE consultancy programme in general.

The significant and positive effect of the QCCE consultancy for the specific items targeted during the consultancy highlights the importance of a sensitive measure to detect specific improvement, particularly in the context of a consultancy programme with individual adaptation and a wide variation in goals. Hence, the development of a measure that is specifically targeted at the QCCE would enable future research to study the effects of consultancy more refined and, moreover, would rule out possible confounders.

The results of our study underscore the fact that improving QCCE is a challenging task. QCCE improved, on average, with 0.65 on a 7-point scale for the ITERS-R items targeted

during the consultancy and with 0.58 for the ECERS-R items (see [Table 2](#)). It should be noted that although the improvement in the experimental group was significant, most average scores on the four subscales and total score of the four subscales increased, but did not reach the adequate to good level.

These moderate outcomes are comparable to the study by Palsha and Wesley (1998) and the QUINCE study (Bryant et al. 2009), which also found an improvement between half and one scale point. A first possible explanation for the modest improvement may be the intensity of the consultancy. A more intensive consultancy programme with additional support may possibly lead to more improvement. In the present set-up, centre directors received three consultations in total, of which the first consultation was mostly introductory and instructive, while the consultancy for improvement was only discussed during the second on-site consultation and the telephone call, which constituted the third consultancy contact. Improving quality of the environment through consultation may be restricted with the current intensity. As mentioned by Zaslow et al. (2010), research into the effect of programme intensity is difficult to find and is highly dependent on the aims of the intervention. In developing the current QCCE consultancy, we aimed for a time-efficient approach considering the costs benefits. We also reasoned that the QCCE list gave direct insight into which points to improve, because it was filled in by the centre directors themselves, who generally manage financial resources and decide about the programme. Although we have no reason to believe that the points of improvement were insufficiently evident for centre directors, effects might have been larger if we would have monitored the process closer by visiting the centre director for a third on-site consultation to monitor implementation and the status of the improvements.

A second explanation for the relatively modest improvement may lie in the nature of the scoring system used with the ITTERS-R and ECERS-R. In the present study, we used the stop-scoring procedure, in which indicators of higher quality are rated only when requirements of indicators of lower quality are met. Gordon et al. (2013) recommend scoring all indicators of items instead of the usual stop-scoring procedure, because scoring all indicators provides centres with more information about improvement points and current strong aspects of quality. It may be possible that centres indeed improved on indicators of the items, but that we were unable to detect this improvement on item level due to the scoring procedure and the stop-scoring procedure may have underestimated the actual improvement made in the care groups. In fact, the study of Hofer (2010) comparing the traditional stop-scoring procedure with scoring all indicators demonstrated that out of 268 care groups, about half ( $n = 135$ ) of the groups improved by one ECERS-R quality category when all indicators were scored.

### **Limitations and future directions**

Our study was not without limitations. First, because the intervention aimed to improve quality of the child care groups, we assigned groups and not centre directors to the experimental or control condition and, therefore, we were unable to compare centre directors in the experimental and control condition. This procedure could have led to diffusion of treatment and we acknowledge this as an important possible threat to the internal validity of our study. To rule out diffusion of treatment in future research, an RCT study design with random assignment at centre director level is warranted.

Second, the QCCE consultancy was conducted parallel to the CIP training for caregivers of the same care groups. Although the CIP training was not directed at improving global QCCE and we controlled for possible effects of the consultancy programme, possible confounding of the consultancy programme and the parallel CIP training for caregivers could not be completely ruled out as an explanation for the effects of the QCCE consultancy. Future research could disentangle the effects of both intervention components by using a full-factorial design with four conditions, namely (1) only QCCE consultancy directed at centre directors; (2) only CIP training directed at caregivers, (3) QCCE consultancy paired with the CIP training, and (4) no intervention at all.

Future research should address several topics. First, future research could focus on new elements of the consultancy (such as focusing on how the centre director conveys the points for improvement to the caregivers of the care group and how the director can monitor these action points) and examine whether these elements enhance effects of the consultancy intervention. Finally, future research could address the optimal dosage for consultancy in child care groups more carefully. Dosage of the present consultancy programme (three sessions) was considerably lower than in the QUINCE study (19.3 visits on average). Results of the present study are nevertheless comparable to the outcomes of the QUINCE study (Bryant et al. 2009). It is important to examine dosage related to effectiveness of the consultancy further in future studies, because of time and cost efficiency of consultancy.

The current study introduces a consultancy programme to enhance QCCE for young children. Results of this first RCT study are promising and comparable to previous consultancy programmes (Palsha and Wesley 1998; Wesley 1994). The outcomes suggest that with self-assessment by the centre director and support from a consultant, centre directors are able to make improvements in the QCCE. Nevertheless, the effects are modest and scrutiny is certainly warranted. Future research should address ways to improve the effects of the consultancy programme. Furthermore, a broader implementation of the consultancy programme (i.e. in different type of groups; including infant and preschool groups, or groups with different quality levels) is needed to examine whether the favourable outcomes of this study can be generalised to other child care settings.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This work was supported by a grant from the Ministerie van Onderwijs, Cultuur en Wetenschap (Dutch Ministry of Education, Culture and Science) to the NCKO (Netherlands Consortium for Research in Child Care).

## References

Belsky, J., D. L. Vandell, M. Burchinal, K. A. Clarke-Stewart, K. McCartney, M. T. Owen, and 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.



- Bloom, P. J., and M. Sheerer. 1992. "The Effect of Leadership Training in Child Care Program Quality." *Early Childhood Research Quarterly* 7: 579–594.
- Bryant, D., P. Wesley, P. Burchinal, J. Sideris, K. Taylor, C. Fenson, and I. Iruka. 2009. *The QUINCE-PFI Study: An Evaluation of a Promising Model for Child Care Provider Training, Final Report*. Chapel Hill, NC: FPG Child Development Institute.
- Campbell, P. H., and S. A. Milbourne. 2005. "Improving the Quality of Infant Toddler Care Through Professional Development." *Topics in Early Childhood Special Education* 25: 3–14. doi:10.1177/02711214050250010101.
- De Kruif, R. E. L., J. M. A. Riksen-Walraven, M. J. J. M. Gevers Deynoot-Schaub, K. O. W. Helmerhorst, L. W. C. Tavecchio, and R. G. Fukkink. 2009. *Pedagogische Kwaliteit van de Opvang Voor 0- tot 4-Jarigen in Nederlandse Kinderdagverblijven in 2008* [Pedagogical Quality of Child Care Centers for 0- to 4-Year-Olds in The Netherlands 2008]. Amsterdam: Netherlands Consortium for Research in Child Care.
- De Kruif, R. E. L., H. J. Vermeer, R. G. Fukkink, J. M. A. Riksen-Walraven, L. W. C. Tavecchio, M. H. Van IJzendoorn, and J. Van Zeijl. 2007. *De Nationale Studie Pedagogische Kwaliteit Kinderopvang: Eindrapport Project 0 en 1* [National Study Pedagogical Quality of Child Care Centers: Final Rapport Project 0 and 1]. Amsterdam: Netherlands Consortium for Research in Child Care.
- Erchul, W. P. 1987. "A Relational Communication Analysis of Control in School Consultation." *Professional School Psychology* 29: 43–51.
- Fiene, R. 2002. "Improving Child Care Quality Through an Infant Caregiver Mentoring Project." *Child & Youth Care Forum* 31: 79–87.
- Fukkink, R. G., and A. Lont. 2007. "Does Training Matter? A Meta-Analysis and Review of Caregiver Training Studies." *Early Childhood Research Quarterly* 22: 294–311. doi:10.1016/j.ecresq.2007.04.005.
- Gevers Deynoot-Schaub, M., R. Fukkink, M. Riksen-Walraven, R. De Kruif, K. Helmerhorst, and L. Tavecchio. 2009. *De NCKO-Kwaliteitsmonitor* [The NCKO Quality Monitor]. Amsterdam: SWP.
- Gordon, R. A., K. Fujimoto, R. Kaestner, S. Korenman, and K. Abner. 2013. "An Assessment of the Validity of the ECERS-R with Implications for Measures of Child Care Quality and Relations to Child Development." *Developmental Psychology* 49: 146–160. doi:10.1037/a0027899.
- Harms, T., and R. Clifford. 1989. *The Family Day Care Rating Scale*. New York, NY: Teachers College Press.
- Harms, T., R. M. Clifford, and D. Cryer. 1998. *Early Childhood Environment Rating Scale-Revised*. New York, NY: Teachers College Press.
- Harms, T., D. Cryer, and R. M. Clifford. 2003. *Infant/Toddler Environment Rating Scale-Revised*. New York, NY: Teachers College Press.
- Helmerhorst, K. O. W., J. M. A. Riksen-Walraven, R. G. Fukkink, L. W. C. Tavecchio, and M. J. J. M. Gevers Deynoot-Schaub. 2017. "Effects of the Caregiver Interaction Profile Training on Caregiver-Child Interactions in Dutch Child Care Centers: A Randomized Controlled Trial." *Child & Youth Care Forum* 46 (3): 413–436. doi:10.1007/s10566-016-9383-9.
- Helmerhorst, K. O. W., J. M. A. Riksen-Walraven, H. J. Vermeer, R. G. Fukkink, and L. W. C. Tavecchio. 2014. "Measuring the 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.
- Helmerhorst, K. O. W., J. M. A. Riksen-Walraven, M. J. J. M. Gevers Deynoot-Schaub, L. W. C. Tavecchio, and R. G. Fukkink. 2015. "Child Care Quality in the Netherlands Over the Years: A Closer Look." *Early Education & Development* 26: 89–105. doi:10.1080/10409289.2014.948784.
- Hofer, K. G. 2010. "How Measurement Characteristics Can Affect ECERS-R Scores and Program Funding." *Contemporary Issues in Early Childhood* 11: 175–191. doi:10.2304/ciec.2010.11.2.175.
- Kontos, S., C. Howes, and E. Galinsky. 1996. "Does Training Make a Difference to Quality in Family Child Care?" *Early Childhood Research Quarterly* 11: 427–445. doi:10.1016/S0885-2006(96)90016-2.

- Layzer, J. I., and B. D. Goodson. 2006. "The 'Quality' of Early Care and Education Settings: Definitional and Measurement Issues." *Evaluation Review* 30: 556–576. doi:10.1177/0193841X06291524.
- Palsha, S. A., and P. W. Wesley. 1998. "Improving Quality in Early Childhood Environments Through On-Site Consultation." *Topics in Early Childhood Special Education* 18: 243–253.
- Schulz, K. F., D. G. Altman, and D. Moher. 2010. "CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomised Trials." *British Medical Journal* 340: c332. doi:10.1136/bmj.c332.
- Snow, C., and S. Van Hemel. 2008. *Early Childhood Assessment: Why, What and How? Report of the Committee on Developmental Outcomes and Assessment for Young Children*. Washington, DC: National Academies Press.
- Vandell, D. L., J. Belsky, M. Burchinal, L. Steinberg, N. Vandergrift, and 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 (3): 737–756. doi:10.1111/j.1467-8624.2010.01431.x.
- Vandell, D. L., and B. Wolfe. 2000. *Child Care Quality: Does It Matter and Does It Need to be Improved?* Madison: Institute for Research on Poverty, University of Wisconsin-Madison.
- Vermeer, H. J., M. H. Van IJzendoorn, R. E. L. De Kruif, R. G. Fukkink, L. W. C. Tavecchio, J. M. A. Riksen-Walraven, and J. Van Zeijl. 2008. "Child Care in The Netherlands: Trends in Quality Over the Years 1995–2005." *The Journal of Genetic Psychology* 169 (4): 360–385.
- Wesley, P. W. 1994. "Providing On-Site Consultation to Promote Quality in Integrated Child Care Programs." *Journal of Early Intervention* 18: 391–402. doi:10.1177/105381519401800408.
- Wesley, P. W., and V. Buysse. 2004. "Consultation as a Framework for Productive Collaboration in Early Intervention." *Journal of Educational and Psychological Consultation* 15: 127–150. doi:10.1207/s1532768xjepc1502\_2.
- Zaslow, M., T. Halle, and K. Tout. 2011. "Differing Purposes for Measuring Quality in Early Childhood Settings: Aligning Purpose with Procedures." In *Quality Measurement in Early Childhood Settings*, edited by M. Zaslow, I. Martinez-Beck, K. Tout, and T. Halle, 389–410. Baltimore, MD: Paul H. Brookes Publishing.
- Zaslow, M., K. Tout, T. Halle, J. V. Vick Whittaker, and B. Lavelle. 2010. *Towards the Identification of Features of Effective Professional Development for Early Childhood Educators*. Washington, DC: U.S. Department of Education.

## Appendix

Table A1. Items targeted during consultancy and numbers of care groups for which the items were targeted (experimental group,  $n = 35$ ).

Item targets	<i>n</i>
Space and furnishings	
Indoor space	2
Furniture for care, play, and learning	17
Furnishings for relaxation and comfort	28
Room arrangement for play	24
Child-related display	34
Space for gross motor play	14
Gross motor equipment	20
Activities	
Fine motor	30
Art	16
Music/movement	34
Blocks	35
Dramatic play	27
Sand/water	18

(Continued)

Continued.

---

Item targets	<i>n</i>
Nature/science	27
Math/number	35
Use of TV, video, and/or computers	3
Promoting acceptance of diversity	33
Language	
Encouraging children to communicate	1
Using language to develop reasoning skills	4
Books and pictures	30
Program structure	
Schedule	3
Free play	7
Group time	1

---